INDIAN AGRICULTURAL
RESEARCH INSTITUTE, NEW DELHI.

54443

THE

QUARTERLY JOURNAL

OF

ECONOMICS

VOLUME LXI

CONTENTS FOR VOLUME LXI

WRITERS

SUBJECTS

PROFITS UNDER NAZI PLANNING[1]

SUMMARY

Early developments: partial planning, 1. — Shift to central planning, 2. — Results, 3. — The rôle of profits, 4. — Supplementary inducements, 5. — Control of investment incomes, 7. — Results, 9. — Control of profit margins: equalization, 10; manipulation, 11; determination, 12. — Results, 14. — The abortive profit ceiling, 16. — Results, 18. — Profit incentives in price control, 19. — Subsidies, 21. — Results, 22. — General conclusions, 24.

It is commonly assumed that central economic planning and profitability of business are incompatible: either comprehensive economic planning without a profit system or business for profit but no central planning must prevail. A study of the Nazi economy, however, reveals that for the first time profitability of business became compatible with central economic planning. This raises two interesting questions: What were the main features of fascist economic planning; how were the principles of both profit and planning modified and integrated effectively in the economy?

Economic Planning

The Nazis practiced extensive partial planning from the fall of 1936 to the middle of 1942, and central planning for the last period of the regime. The government had introduced central planning at the beginning of the war, through the decrees of August and September, 1939. Consumption and income of civilians were reduced; the working week was generally raised to 60 hours; compensation for "overtime" and other bonuses had to be paid as a tax to the treasury. But the quick success in the Polish campaign

1. The present paper is part of a research project on government and business under the Nazis conducted by the School of Business, University of Chicago.

led to a repeal or modification of the original plans for total economic mobilization. The plans for concentrating production in large plants were shelved; small and middle-sized plants were incorporated in the military sector through dispersal of orders.[2] In fact, partial mobilization replaced the original blueprints for central planning and total mobilization.

" Partial planning was predicated upon "lightning" military campaigns. On the one hand, this form of warfare minimized the military demands upon the economy. The short periods of attack limited both the material used by the conquering forces and the destruction of properties by the opponents in the campaign. This form of warfare, deliberately devised by the Nazis, fitted the situation of Germany with her limited resources, and opened the opportunity for maximum gains in war booty. The military superiority of the German armed forces could be exploited without exhausting their own economy; victory could be secured without devastating the economies of conquered lands, destined to be a part of the fascist empire. On the other hand, the periods of lull set off the various campaigns from each other, defined the specific tasks for planning, and provided a valuable period of economic preparation for the campaign. The reduced scope of military requirements and well-defined tasks for war production made economic planning predictable and manageable. In consequence, the specific economic planning for brief campaigns dispensed with the need for an overall economic plan of many years duration, and planning for war was reduced to the proportions of a series of partial plans.

But when Allied military action imposed a war of attrition upon Germany, the Nazis had to shift to total economic mobilization for an all-out war. Central planning was the means by which they hoped to realize this goal. Relative to the previous economic performance, total mobilization was a success. The index of finished armament production rose from 100 in 1942 to 322 in July, 1944.[3] The number of aircraft produced yearly increased from 8,200 in 1939 to almost 40,000 in 1944. Simultaneously, consumption fell considerably. Civilian consumption of

2. Georg Wolff, "Mittel- und Kleinbetriebe im Kriege," *Wirtschaftskurve*, February, 1940.

3. Rolf Wagenfuhr, *Deutschlands Kriegswirtschaft, 1939–45*, captured manuscript, quoted in U. S. Strategic Bombing Survey, The Effects of Strategic Bombing on the German War Economy, Washington, 1945.

leather and leather substitutes for shoe manufacture and repair declined by 55 per cent between 1938–39 and 1943–44. Of the shoes still produced in 1944, civilians received only 28 per cent, as compared with 64 per cent in 1938–39. In January, 1945, the daily ration for "normal" consumers had a total content of 1,721 calories. This was almost 300 units below the minimum emergency level of consumption, and only a little above the ration set by the American Military Government after the occupation. The expansion of the military and reduction of the civilian economy are the more remarkable because Germany had neither idle facilities nor large consumer goods industries as a source of war production. Total mobilization was achieved on the basis of full employment. It was secured through the extensive utilization of the resources of the Empire and reduction of the civilian economy.

How effective was central planning under the Nazis? It is too early to tell the whole story and examine its effects, but three facts are beyond doubt.[4] First, central planning came two years too late. Considering the major task of total mobilization for defeating Russia before the United States could become an effective belligerent, central planning should have been instituted in 1940, instead of 1942. Second, central planning was coupled with the task of a partial reorganization of business. The available resources and facilities for production of armaments had to be increased through an extensive rationalization and efficiency drive within business units, through a shift of labor and machinery to the most efficient plants and the laying-off of others, and through reduced production of civilian goods. The fairly large success of these efforts increased the available resources, but also changed continually the conditions for central planning. Third, the realization of plans was increasingly disturbed and upset by Allied air attacks. Endangered or bombed plants called for dispersal of essential factories. Bombing of cities led to the uneconomical use of labor in the clearing and repair of bombed areas. Bombardment also limited the extent to which consumption could be reduced because of the losses of consumer goods by air attacks. In consequence of the last two factors, central planning had to take place under a high degree of

4. The shift from partial to central planning did not initiate a change in the power relationships between government and business. The party continued to be the major, big business the minor, holders of power. Cf. Arthur Schweitzer, "Big Business and the Nazi Party in Germany," *Journal of Business*, January, 1946.

uncertainty and be executed under the most unfavorable conditions.

Present information, however, does not justify the conclusion that central planning was one of the major reasons for the collapse of the German economy. Instead, two causes led to the quick disintegration of the economy: the loss of the resources of the empire; the increasing effectiveness and devastation of the Allied air attack.

What was the rôle of economic planning in the Nazi economy? Clearly, it was one of the new features introduced by the Nazis into the German economy. Reduced to its essentials, it involved basic changes in the traditional economy in important respects. First, economic freedom was replaced by the principle of universal obligation to work in accordance with governmental instructions (Arbeits- und Wirtschaftspflicht). Second, trading in private markets was replaced by comprehensive governmental regulation of almost all markets in the economy. Third, freedom of association was displaced in favor of compulsory business organizations and effective operation of governmental agencies in business. Finally, the function of unconscious direction of the economy through the profit motive was taken over by economic planning of activities by the government.

THE RÔLE OF PROFITS

The profit situation after 1936 was characterized by two strange paradoxes. On the one hand, the Nazis gradually suppressed the profit motive but retained profit as a stimulus for producers to increase output and achieve efficiency. On the other hand, government price control determined the permissible profit margins in fixed prices but still offered opportunities for businessmen to realize substantial total profits each year. These contradictions produced two problems. How and why did the Nazis sever the profit motive from the profit stimulus? What was the effect of extensive governmental price fixing upon the rôle and use of realized profits in the economy?

In private capitalism the profit motive is a powerful force. In the sphere of production it performs at least two functions. The one affects the actions of producers; the other influences the general course of the economy. Actual or expected profit opportunities constitute incentives for owners to reinvest their capital,

increase output, improve efficiency of production. These incentives enlarge the opportunities and modify the prevailing conditions within operative plants. In addition the profit motive affects the relationship between firms and industries: actual or expected differences in the rate of profit between firms and industries induce owners to redirect their capital from less to more profitable fields of production. This shifting of capital and employment has an effect not consciously sought; it determines the total volume of business and directs the general course of economic activities.

The Nazis refused to accept these traditional functions of the profit motive. On the contrary, they effectively separated its two functions in production. Profit as a stimulus for private owners was retained but greatly modified. Profit as a guide for the general direction of the economy was suppressed, and its place was taken by the economic plans of the government.

Governmental planning set the prime goals for the economic activities of persons and firms. Governmental agencies, implementing plans, allocated factors of production among concerns and industries. Commissions and inspectors reduced the civilian industries in favor of war industries. Innumerable agencies regulated the exchange of all products, services, and almost all property titles. The general direction of the economy was no longer left to the profit motive. The course of economic activity was planned and directed by governmental agencies.

What happened to profits as a stimulus for producers? As usual in major decisions, a conflict arose between party leaders and big business. Leaders of big business favored the full operation of the profit motive. Party leaders desired intensified production, but insisted upon extensive control of all forms of profits. Under the increasing pressure for efficiency, neither of the rivals carried his point. The resultant compromise led to the incorporation of profit and other supplementary inducements into a system of profit fixing.

The two outstanding supplementary inducements introduced by the Nazis were the contests among enterprises for the title of a "model firm" and payments of prizes and bonuses for special achievements in production. The idea of a national competition among firms for the title of a model firm originated in the Labor Front. It was the means by which this party organization induced employers and employees to accept, and operate through, its

division of "Strength through Joy." The winning firms were selected according to their ability to (a) maintain social peace, (b) provide welfare schemes for their employees, (c) increase the productivity of labor, (d) and follow closely the Nazi economic policies. During the war, however, the last two criteria alone became relevant. Most of the welfare schemes were discontinued. Maintaining social peace became a responsibility of the Labor Front and the Secret Police. Of equal significance in selection of contesting firms became the ability of plants to utilize raw materials and equipment efficiently. In 1941 a total of 290,322 enterprises participated in the contest; 119 were chosen as model firms; 2,000 received regional diplomas; and 3,583 obtained rewards of achievements for more than average efficiency.[5]

The bonus system fostered a variety of activities. Suggestions for improvement of production methods or discoveries of new processes brought many a foreman and manager a small bonus. Replacement of scarce by more plentiful materials carried governmental premium payments (e.g. timber bonus). Prizes were paid to firms which introduced new labor-saving machines. During the war, however, cash payments did not incite many to special achievements. Hence premiums were granted in kind. Firms that utilized coal sparingly were permitted to store a two-months supply for their own use. Plants conserving scrap were given a preference for speedy delivery of allocated materials. Efficient management was a recommendation for being granted a lease of a sequestered enterprise in occupied lands. For effective managers were at a premium and special management schools tried in vain to satisfy the most urgent demand.

The special inducements had a limited success. The national contests among enterprises, at first a mere political propaganda stunt, became general drives for increased output, improved efficiency, and conservation of resources. The winning firms increased their chances of receiving the favorable attention of regulating governmental agencies. Such special consideration became increasingly important when the contests were followed by drives to close the inefficient plants and shift the skilled laborers from the less to the more efficient concerns. Prizes and premiums had a similar effect for the winners. For premiums did not conflict

5. "Leistungskampf der deutschen Betriebe," *Soziale Praxis*, 1941, p. 549.

with government plans. They became recognized devices to break bottlenecks. Prizes stimulated firms to special achievements in fields in which accomplishment threatened to fall short of government plans. Consequently, contests and premiums were incorporated in the production and rationalization plans of the government.

However effective in implementing plans, the special inducements did not replace the profit stimulus for producers, as some Nazis desired. Fixing of profits, not their suppression, was the official policy of the Nazi party. Government profit fixing tried to impose three different lines of action upon owners. First, profits usually spent in consumption were reduced in favor of profits invested in governmentally desired projects. Second, the rates of profit between firms and industries were manipulated in order to equalize the rate of profit among desired activities and discriminate against the primarily civilian industries. Third, a significant attempt was made to establish a general profit ceiling that was to supplement the policy of fixing profit margins through price control. How far were these profit policies of the Nazis successful?

CONTROL OF INVESTMENT INCOMES

The Nazi government effectively discriminated against the rentiers, who usually spend a relatively large portion of their property income for consumption. Dividends, interests, and rents were deliberately depressed. The various rates of interest were greatly reduced. The yield of long-term government bonds declined from 7 per cent in 1933 to 4.5 per cent in 1939. Rising bank deposits reduced the yield further to 3 per cent in 1942. The discount rates of the central bank declined from 8 per cent in 1932 to 2⅛ in 1941. The short-term rates of the private banks fell correspondingly. The result was a marked decline in interest income. The property holders who lived mainly from interest income protested[6] but their complaints were disregarded. A rapidly rising public debt in the face of drastically reduced rates of interest testified to the effectiveness of governmental control of money and capital markets.

Similarly, shareholders were not permitted to benefit fully from the rising net income of corporations. The dividend limitation law of 1934 compelled corporations to pay any excess over 6 (or 8)

6. Cf. "Eine Neue Zinsetappe," Bank-Archiv, January 15, 1941.

per cent to a government fund. Excess dividends were invested in government bonds in the name of the shareholders. Yet this fund remained small: total payments from 1934 to 1940 did not exceed 108 million marks. A portion of these accumulated dividends was disbursed in the form of non-interest bearing certificates; the remainder was kept by the government for "the duration of the war." Although not many "excess" dividends were paid to the government, the law of 1934 set an effective upper limit upon dividend payments. For the actual rate of dividends of corporations rose from 2.9 per cent in 1932–33 to 5.6 per cent in 1938–39. Hence most of the dividends paid remained below the maximum set by the government.

Lower rates of interest and limited dividends at a time of rising net profits produced a marked increase in the undistributed profits of corporations. Official sources estimated that undistributed profits amounted to only 170 million marks in 1933. They rose to 1,200 million marks in 1935 and 3,420 million marks in 1938. This remarkable advance exceeded by far the actual dividend payments, which amounted to 1,200 million marks in 1938. In other words, undistributed profits were nearly three times as large as the dividend payments to shareholders.[7] The Nazis welcomed and intensified the corporations' policy of keeping the major portion of net profits in the business. The government closed the capital market for many years to corporations, compelling them to rely upon their own resources. Government armament orders and new projects called for expansion and investment of capital. The financing of such projects was to a large extent accomplished by ploughing profits back in the business. As a result capital formation increased rapidly at the expense of the consumptive portion in property income.

Yet the Nazis had to modify their policy of holding dividend payments down. The boom of the "free" security market called for a greater supply of shares. Corporations were permitted to issue more shares on the basis of their increased assets. But not many availed themselves of this opportunity. The 10 per cent capital-issue tax was an effective deterrent. Instead of lowering this tax, the Nazis imposed a dividend tax in 1941 which repealed

7. In addition, undistributed profits of 1,000 million marks were made by individual proprietorships and partnerships. This increased the total of undistributed profits to 4,420 million marks in 1938. Cf. "Selbstfinanzierung und Kapitalmarkt," Bank-Archiv, 1941, p. 174 ff.

the dividend limitation law of 1934. The revenue from the new dividend tax was of small import, however. The law offered corporations three alternative lines of action: to write up their capital, build up undistributed profits as in the past, or pay the new tax. The majority of the corporations accepted the first two alternatives. About 30 per cent of the joint-stock companies translated some of their reserves into capital stock. These companies represented 59 per cent of the total stock.[8] Most of the "baby shares" were issued to the existing stockholders without payment. But dividends rose merely by 13 million marks, which was a small portion of the total dividends received. Hence increase in capital stock did not reverse the policy of discrimination against prospective consumers of property income.

What lessons can be drawn from this line of profit policy? In the first place, the Nazis refrained from a fundamental change in the formation of property income in Greater Germany. Capital invested was recognized as the major source of property income. The connection between the private property structure and the formation of income from property, typical for private capitalism, prevailed throughout the Nazi regime. In the second place, the Nazis interfered deliberately in the distribution of realized property income among the various groups of owners. Whoever was suspected of consuming a share of his rent, interest or dividend income had to face a persistent governmental discrimination. In fact, the government established two different groups: the "esteemed" and "despised" receivers of property income. The former group of active producers and investors was given the opportunity to transfer a share of income to themselves which otherwise would have gone to the rentiers.[9] But the federal treasury also benefited from this governmentally sponsored income transfer between owners, for the extraordinary fall in the rates of interest reduced the burden of servicing the public debt.

In the third place, the goal of the income transfer was to

8. H. W. Singer, "Recent Conditions in Germany," Bulletin, London and Cambridge Economic Service, 1943, p. 88.

9. In effect, this governmental discrimination favored the upper class and hurt the middle class, because the largest producers and investors alone got the full benefit of the income transfer. Most of the rentiers, however, belonged to the middle class. Hence the discrimination between income receivers corresponded to the separation of owners into desirable and undesirable property-holders. Cf. Arthur Schweitzer, "Big Business and Private Property under the Nazis," Journal of Business, April, 1946.

increase the formation of capital. The "esteemed" group of income receivers had to reinvest its capital in governmentally acceptable projects. The lines of investment became increasingly determined by the government. Indirect control was exerted through placement of government orders for materials and plants, as well as through suppressing demand of civilian producers and consumers. Direct control of investment became the function of the government agencies that allocated materials and manpower. Requests for industrial construction permits, for instance, were scrutinized more effectively with each additional year of war. This is evident from the decline in the physical volume of total construction, which fell from 13 billion in 1938 to 3.8 billion marks in 1944.[1] Finally, the functions involved in the formation and reinvestment of capital were divided between big business and government agencies. Private concerns were usually limited to direct investment of undistributed profits. Sale of industrial bonds was permissible only for a few large concerns. Hence the banks almost lost their customers in the industrial field. The funds of banks and insurance companies had to be invested in government bonds and bills. When the war ended, about 80 per cent of the assets of banks and insurance companies were invested in government securities. Indirect investment had developed into a governmental monopoly. This was exercised by the central bank, which regulated the capital and money markets. It was through the effective control of these markets that the government increased the national debt from 37.3 billion marks in August, 1939, to 400 billion marks in June, 1945,[2] without a decisive effect on price levels. For indebtedness was the preferred method of Nazi war finance, confiscatory taxes and expropriation being primarily practiced in the Nazis' empire.

CONTROL OF PROFIT MARGINS

How did the price commissioner handle profits through price control? Four lines of actions can be discerned. These involved freezing, reducing, manipulating and determining of profits. Actually, freezing of profits at the prevailing level of September, 1936, was merely a preliminary step: it laid the foundation for an

1. Calculated in prices of 1938 and given in "Die Deutsche Industrie," a captured manuscript, quoted in The Effect of Strategic Bombing on the German War Economy, p. 55.
2. Economist (London), November 20, 1943, and September 29, 1945.

extensive determination of prices and profit margins. Reduction
of profits was accompanied by governmental subsidies. Both
became effective means of equalizing the rate of profits between
firms and industries. In consequence, the profit policies of the
price commissioner can be discussed under three headings: equal-
ization, manipulation and determination of profits.

Equalization of the rate of profits began with government
subsidies as well as with reduction of profits by government actions.
Industries producing substitutes or military equipment could and
did receive subsidies. Most of them were granted indirectly,
either through favorable prices or long-term purchasing contracts
concluded between producers and government agencies. Cash
subsidies were paid to producers to prevent losses and assure
reasonable profits. The various forms of government subsidies
fulfilled two functions. On the one hand, costs of experimentation
in new fields and products were shifted to the government through
subsidy payments; on the other, rising costs of production in many
cases did not lead to higher prices, the increasing costs being
either taken over by the government or shifted to more favorable
producers. The groups suffering most under increasing costs were
farmers, as well as coal and iron producers. During the war,
government subsidies rose tremendously. Premiums paid by the
state were estimated at 1000 million marks annually.[3] This
amounted to nearly 10 per cent of the total receipts from farm
sales. Rising costs in the coal mining industry were compensated
in 1940 and 1941 by increased export prices. When this was no
longer possible, the coal concerns received a government subsidy.
The British military government· in 1945 found a coal price of
18 marks per ton for the home market, whereas the cost of pro-
duction had risen to 50 marks per ton. Hence the longer the war
lasted, the more did productivity of labor decline and subsidies
rise. No wonder the Nazis tried all possible devices to shift a
part of the increasing cost to other producers who were more
favorably situated. This gave rise to a deliberate policy of equaliz-
ing the rate of profits between firms and industries.

Inter-business payments originated in the fields of agriculture
and imports. As early as 1934, producers of milk for consumers
compensated producers of processed milk for their low prices.

3. "German Europe: Controlling Prices," *Economist*, October 31,
1942, p. 545.

Importers had to pass on a portion of their profits to home producers. In 1935 the well-known "export tax" was introduced, through which the home industries paid a part of their profits to governmentally recognized exporters. With the beginning of 1937 sugar refineries compensated beet producers, breweries compensated barley producers, wheat flour mills passed on a part of their profits to rye flour mills. Costs of transportation were shifted from producers to dealers. Industries with high profits were compelled to reduce their prices. For instance, chemical firms had to reduce the price of fertilizers repeatedly so as to help the farmers. Producers of marked-up products had to reduce prices in order to increase the profit margin of dealers and lower the prices for consumers. Similarly, vertical combines were urged to equalize profits within their spheres of influence. For instance, in the iron and steel industry prices of products were kept almost stable, though the cost of production of crude steel had risen greatly. The steel combines made great profits in steel products, but lost heavily in crude steel.[4] Under the direction of the price commissioner, steel producers had to use excess profits in some plants to offset their losses in others. Incidentally, this intensified the prevailing tendency to vertical combinations. Consequently, rising costs were absorbed through equalization payments which became an important feature of price fixing. Such payments could be made within trusts or combines, between firms in related branches, or between whole industries. Payments could appear as a reduction in prices, a government tax, a surrender of profits, or a change in accounts between subsidiaries in a combine. Whatever their form, loss of profits for one business became a "subsidy" for another business. Through these inter-business payments, the price commissioner equalized the rate of profits between firms and industries. Although limited in scope, equalization payments prevented the more striking cases of an unequal distribution of profits that resulted from governmental planning and direction of the economy. But the price commissioner reached the limit of this policy in the first year of the war: the need for cost absorption exceeded the opportunities of equalization payments between concerns. Hence the government had to increase its subsidies.

Quantitatively more important was the determination of

4. "Die Ertragslage der Schwerindustrie," *Wirtschaftskurve*, May, 1940.

profit margins through government price fixing. Increases in prices after the price-stop decree of September, 1936, induced the price commissioner to handle profit margins in three different ways. Requests for increased prices were refused, if the profit margins of firms were relatively high. Hence high profits became an absorber of increasing costs. But when the rise in cost was continuous, as in the case of imported goods, or when profits could not absorb increasing costs, the commissioner pursued a policy of stable profit margins. Increased cost could be passed on to buyers, but the profit margin had to be kept unchanged. This abolished the traditional practice of calculating profit margins as a standard percentage of the price. Stable profit margins even under conditions of rising prices prevailed for a considerable section of the economy prior to 1941.

Another rule related to products not formerly produced in bulk. This was a common condition in the substitute and in most armament industries. A modified cost-plus-profit price was introduced, first in construction and electricity and then extended to all firms filling orders for the government. This cost-plus-profit price was actually a governmentally calculated price.[5] Two rules were applied by government agencies. A list of items, called "necessary cost," could alone be included as overhead cost in the price; all other items were rejected. The most important items rejected related to inefficiency or prices of raw materials and wages paid above those fixed by the government. Moreover, the recognized profit margin had to be composed of two elements: a "wage" for the entrepreneur and interest for the total capital invested. The rate of interest acceptable was not to exceed the average effective rate of long-term government bonds.[6] When the prices were calculated in 1938, this rate was 4.5 per cent. These two rules had several beneficial effects. Cost-plus-profit prices were no longer a means of squeezing the treasury. Instead, two forms of incentives were built into the calculated prices. Any producer whose actual cost was below the "necessary" cost could keep the difference. Similarly, a producer who paid less than 4.5 per cent for capital borrowed realized a gain. These factors amounted to an incentive for increased efficiency. Yet this was coupled with an

5. Leonhard Miksch, "Vom Preisstopp zur Kostenkontrolle," *Wirtschaftskurve*, August, 1939.

6. Leonhard Miksch, "Der Zins als Kostenfaktor," *Bank-Archiv*, December, 1938.

attempt to make producers a party in wage and price control. Whoever filled government orders and paid more than the official prices and wages had to bear the difference himself. This was a fairly effective device, introduced in September, 1939, to check the circumvention of wage and price controls and integrate the various forms of government planning and direction.

The limitation of the cost-plus-profit price related to the spread of cost between firms in an industry. The government fixed a uniform price on the basis of average "necessary" cost in the industry. Essential producers with higher than average cost could receive a subsidy, whereas non-essential producers were urged to shift to other products within their capacity. Producers with less than average cost realized a "differential" profit. With the beginning of the war, the Nazis objected to such profits. The war economy decree of September 4, 1939, stated the principle that nobody should profit from the war. In one of his speeches Hitler threatened to behead all war profiteers. The demand for skimming off war profits became general, just as laborers were made to forego their overtime payments.[7] Hence the ground was laid for a shift to an overall fixing of total profits realized by firms.

On the whole, fixing of profit margins through price control was fairly successful. Three lessons emerge from this experience. First, the policy of permitting necessary price changes while stabilizing profit margins severed profits from their close association with variations in prices. Profits ceased to be a product of prices. Government policy transformed profits into either a standardized income of a base year or a minimum income of capital invested. These actions put a floor under the profit margins of producers. They also eliminated the risk that price declines might reduce or eliminate profit margins. Yet under conditions of full employment the problem of the price commissioner was how to prevent producers from calculating profits as a standard percentage of prices. Strangely enough, the price commissioner utilized a peculiar tax device for his purposes. The treasury had compelled businessmen in 1935 to keep books on all incoming and outgoing transactions. These books provided the evidence as to how much turnover tax a firm had to pay to the treasury. Price enforcing officers consulted these books on transactions. They

7. "Die Abschöpfung kriegswirtschaftlicher Differentialgewinne," Bank-Archiv, November, 1939, p. 526.

compared the books of firms, checked upon the actual prices charged, and thereby determined the profit margins received by sellers. Hence the danger of discovery was great and most of the firms accepted the stable profit margins in their price calculations. Yet fixing profit margins did not limit the total profits realized by producers in a year. It merely determined the profit per unit sold. With the beginning of the war the Nazis became concerned over the total profits of producers.

A similar success was scored, secondly, relative to the cost-plus-profit prices. The Nazis licked two difficulties of this method of pricing. The standardized cost accounting system, imposed upon business in 1937, gave price offices full opportunity to check upon the actual costs of production. The concept of "necessary cost" was based upon actual cost; both enabled the Nazis to eliminate the producers' practice of including simulated cost in their prices. Similarly, the "plus profit" item was replaced by governmentally calculated profits as a percentage of capital invested, instead of the cost of production. Two other difficulties remained, however: the spread of costs between producers, and differences in the degree of utilizing capital. In averaging the "necessary" cost of production and granting an average profit on capital invested, the Nazis permitted the most efficient producers to realize a special profit. Hence it was excess profit, resulting from efficiency, not from simulated costs of production, that induced many Nazis to attack this method of pricing war materials.

In the third place, the attempts to equalize the rate of profits between fields of investment were widely opposed by businessmen and regarded with misgivings by Nazis. Clearly, comprehensive direction of markets and investment by the government had stifled the inherent tendency towards equalization of profit rates. On the contrary, many a governmental action created or intensified disparities in the rate of profits realized by producers. The natural thing for the Nazis to do was to make equalization of profits a function of the government, but the two methods selected for this policy were not quite satisfactory to the Nazis. To be sure, cost absorption was enforced through the imposed cost accounting and pricing systems; yet under conditions of rising costs the profit opportunities declined and at the beginning of the war the Nazis were still reluctant to pay the price demanded for cost absorption by big business: penetration into the fields of small business and

increased profits through vertical combination. Equally, the Nazis administered inter-business payments successfully through equalization funds. Rising costs, however, limited the amounts of profits available for payments. The Nazis did not conclude that equalization was impossible but that better methods must be found to pursue the same policy. For they were determined to reduce government subsidies by shifting more profits between producers.

THE ABORTIVE PROFIT CEILING

The new policy adopted aimed at a fixing of the total profits realized by a firm, which was to be added to the fixing of profit margins for individual products sold. The new policy was to avoid the excess profits accruing to the most efficient producers in cost-plus pricing, and strengthen the policy of equalizing the rate of profits between fields of investment. The new policy was called "profit stop"; actually, it introduced a form of ceiling upon yearly profits of concerns.

Ceilings upon total yearly profits were preceded by two new theories. One gave the reason, the other tried to justify the government ceiling on total profits. The crucial point, we are told,[8] in price and production policies is the effective direction and reduction of demand. Rationing limits the demand for consumers' goods; allocation limits and directs the demand for producers' goods. Government permits introduce a new system of distribution for goods: effective demand is determined by permits as well as by the income and desires of buyers. Permits reduce total demand beyond the limits set by desires and income. In effect, permits nullify the impact of excess purchasing power upon demand and prices. Hence permits adjust demand to the reduced supply and thereby provide the basis for effective price control. Any war, however, creates a tendency to increasing costs of production. This is the result of various factors: use of substitute materials, utilizing less productive workers, assigning orders to less efficient plants, producing new kinds of products, and additional overhead cost for organizational adjustments. Higher costs of production exert pressure upon fixed prices. The price commissioner will be called upon to permit rises in prices. The author expected in the

8. Leonhard Miksch, "Mark und Marken," `Wirtschaftskurve, May, 1940.

spring of 1940 that the rise in cost would be limited; the pressure would not constitute a serious threat to the "stable" price level. A year later, however, the price commissioner was faced with a tremendous pressure upon prevailing prices and a marked deterioration in the quality of goods. He was searching for a device to offset the impact of increasing costs upon prices.

The second theory gave a new interpretation of the sources of profits. It distinguished between two kinds of profits according to their origin.[9] Profit of superior performance (Leistungsgewinn) results from great efficiency of enterprises. Differential profit (Differentialgewinn), however, is the product of successful speculation, advantages of location or transportation, wage differentials, or monopolistic agreements among producers. These two kinds of profits are not of equal significance in the economy. Profits of superior performance are justified, because they spring from managerial ability, organizational improvements, and technological advances. These achievements should receive compensation in the form of reasonable profits; but the same cannot be said for differential profits, because owners reap products in an area which they have never cultivated. Hence differential profits should either be paid to the treasury or be utilized in plants for essential investments which have no profit expectations. This profit theory, in effect, suggests that only managerial ability and plant efficiency should furnish a justification for profits, not invested capital in its manifold forms.

The price commissioner did not quite follow this theory when he issued his "profit stop" decree in March, 1941. He distinguished between permissible and excessive profits. The latter had to be surrendered. Permissible profits had to be equal to the profits of a normal year prior to the war. For industrial concerns permissible profit was influenced also by turnover. If capital was turned over once a year, the permissible profit before taxes should not exceed 4.5 per cent of the capital employed. To this could be added 1.5 to 3.5 per cent as a risk premium and 1 to 2.5 per cent as a wage for entrepreneurs.[1] Although larger industrialists were given a chance to negotiate their permissible profit with the price commissioner, his intention was to set a ceiling upon the rate of profits and thereby

9. Ludwig Preller, "Differentialgewinn," *Soziale Praxis*, 1941, p. 91 ff.
1. For details see Ernest Doblin, "The German 'Profit Stop' of 1941," *Social Research*, September, 1942.

limit the total amount of profits available to owners. A ceiling upon profits was not an end in itself, however; it was utilized as a means of easing the pressure upon fixed prices[2] and reducing the volume of government subsidies. Producers had to pay the excess profits of 1940 as a tax to the price commissioner, but excess profits of 1941 had to be translated into lower prices.

The profit ceiling policy failed. The price commissioner, an old *Gauleiter* of the party, fell in disgrace and was dismissed by Hitler. Three repercussions combined to cause his downfall. Administratively, the new method of regulating profits involved extensive red tape. Firms had to (a) compute their profits according to the new formulae, (b) calculate their prices anew, (c) recompute their taxes for 1940, and (d) file a return on profits and prices with the commissioner six months after the end of the business year. This additional administrative burden came at a time when office help was hard to get. Economically, the computation of permissible profits as a percentage of capital invested eliminated the link between output and profits. The volume of goods produced, in spite of the reference to the turnover of capital, was no longer closely connected with the total profits obtainable. Hence producers lost interest in speed of production, lowering of cost, and efficient organization of plants.[3] They could secure their "permissible" profit by less than exceptional efforts. Hence the profit ceiling reduced the effectiveness of the profit stimulus. The price commissioner, in his economic naïveté, had sinned against one major tenet of the Nazi economy: to get the highest total output from the available resources. Politically, the price commissioner had incurred the opposition of other party leaders. He began to collect quasi-taxes, gave directives to government agencies placing orders, and tried to establish the requirements of price fixing as the superior goal of the economy. Some Nazi top leaders therefore supported the opposition of big business against the price commissioner. By obtaining the political support of some powerful Nazis, big business was able to get the policy of profit

2. In March of 1941, when the "profit stop" decree was issued, not less than 3,975 shopkeepers in Berlin were convicted for violating the price regulations. Cf. H. W. Singer, "The German War Economy in the Light of Economic Periodicals," Economic Journal, 1942, p. 26.

3. Some periodicals complained of a reckless inflation of cost. See Singer, ibid., p. 28.

ceiling removed. The price commissioner was dismissed.[4] The profit stimulus was restored: the Nazis recognized it as the most effective inducement to maximum armament production.

PROFIT INCENTIVES IN PRICE CONTROL

Dr. Fischboeck now became price commissioner. He was a former Austrian banker and cartel man, who was well-known to the German steel industrialists. The new commissioner acted decisively: price control was overhauled; profit taxation was severed from price control and transferred to the treasury.

The new excess profit tax cannot stand comparison with the corresponding tax in the United States. Tax base, rate, exemptions and deductions of the tax of April, 1942, were all favorable for producers. Concerns with net income of less than 30,000 marks were exempted. This limited the tax to 30,000 of about 290,000 firms. "Excess profits" were defined as earnings exceeding 50 per cent of the profits in 1938 (or the average of 1936–38). The tax rate was 25 per cent for private firms and 30 per cent for corporations. Not all the taxable profits, however, had to be paid to the treasury. Deductions were granted to firms who invested portions of the tax in desirable projects, reduced prices, or spent it for experimental work. Although the generous exemption and base were reduced somewhat in 1943, it is unlikely that the excess profit tax netted the treasury substantial revenue.

The overhauling of price control did away with the idea of price stability. The new goal of the price law of 1942 was to facilitate the production of essential materials. Hence three new principles were introduced into price control. Multiple prices for the same product sought to overcome the spread of cost among producers; prices were permitted to rise with increasing cost; incentives were built into the system of fixed prices so as to increase production.

The system of multiple prices, introduced in 1942, was limited to armament products. It replaced the cost-plus-profit prices, because the comprehensive cost studies were · said to be too cumbersome. Multiple prices were based upon different levels of efficiency in an industry. This notion grew out of the experience

4. The eclipse of Price Commissioner Wagner coincided with the rise of Gauleiter Sauckel, who took over effective control of the labor market with the help of the Labor Front.

of cost prices. Already the profit stop decree had authorized the price commissioner to allow higher "permissible" profits, if they were a result of efficient management; but this permission remained on paper, because firms were not then classified into efficiency groups. Such groupings of firms came with the price law of 1942. Gradually five different price groups were set up, which covered in 1943 about 70 per cent of all armament products. Prices were fixed by government agencies, not negotiated with firms. Producers, however, could elect their price group. The choice, of course, was influenced by two kinds of government action: inducements and punitive cost investigations. When firms selected the lowest price, they were relieved from the excess profit tax. A specific profit incentive for efficient producers was thus built into the system of multiple prices. When firms refused to elect a low price, government inspectors checked plants, discovered sources of waste, and caused appreciable reduction in cost. The result was a tendency for producers to shift to the lowest possible price group.

Multiple prices enabled the government to achieve two goals. In shifting from average cost to multiple efficiency prices, the government was able to reduce the prices for war materials purchased. The lower prices of the most efficient firms, however, shifted a portion of the "differential profits" from the producers to the government. Moreover, firms with high prices invited government inspection and subsequent cost reduction. Hence fixed multiple prices facilitated the drive for "rationalization" of less efficient firms, and thereby contributed to the paramount goal of maximizing armament production.

In the field of civilian products, however, prices continued to be based upon extensive cost studies. A trend towards an increasing variation in cost between firms was discovered. This produced changes in permitted prices. When costs had increased, the price commissioner allowed a rise in prices. When cost had been reduced through "rationalization," the commissioner ordered a lowering of prices. His aim was to offset the rise by an imposed decline in other prices and thereby nullify the net impact of price changes upon the price level.[5] Cartels and groups were ordered early in 1943 to pre-

5. This policy was accompanied by standardization of products, which were required to have the same quality. Cf. "Umschau," Frankfurter Zeitung, October 10, 1942. Freezing of quality standards was limited to a few industries, however, and lower qualities of civilian goods became the general practice in the last years.

sent proposals for price reductions. The commissioner's order to reduce prices of individual firms was widely publicized. Piece rates of wages were cut in 1943 by government order; the reduced wage cost had to be reflected fully in lower prices. Yet all these efforts did not keep the price level stable. Even the official cost-of-living index rose during the war by 14 per cent.[6] Prices of clothes increased by 38 per cent in the official index. In spite of all punitive and propagandistic actions of party and government, people increasingly lost confidence in the value of money as prices rose.

The partial rise in civilian prices induced the government to grant more premiums to producers. Payments were given primarily to farmers, and were usually made dependent upon some additional actions of producers: to cut down their self-consumption and deliver more to the regular channels of the market, or to produce more goods of a specific kind on the available acreage. These specific premiums were accompanied by general premiums only if adverse weather affected all producers alike. The payments led to a rapidly rising rate of government subsidies, and a remarkable discrepancy between producer and consumer prices in food, coal and steel. The result was a peculiar kind of dual price system in the field of civilian products. The multiple prices in the military and the dual prices in the civilian sectors exerted a different influence upon the price level. In the military field, the government was the sole purchaser and effective cost inspector. It had complete control over demand and supply of the products and could compel producers to sell at the lowest price. In the civilian field, however, there were many producers and consumers. Inspection and reduction of cost was here a slow process. Consumers wanted goods and adopted the well-known attitude of considering the price commissioner as their enemy. Gradually, neither supply nor demand of civilian raw materials and products could be fully controlled by the government. Hence, even before military defeat, price control failed partially in the civilian sector of the economy.

Two factors were responsible for the partial failure of price fixing. On the one hand, the supply of civilian products and related raw materials declined rapidly on account of the loss of the empire and increasing air attacks. On the other hand, the Nazis reëstablished the connection between prices, volume of production, and

6. H. W. Singer, "Recent Conditions in Germany," London and Cambridge Economic Service, October, 1944.

the profit stimulus. Profits of essential producers ceased to be the absorber of increasing cost.[7] Additional items of cost led either to higher prices or to rising government subsidies. Similarly, a larger volume of products and lower "overhead" or subsidies caused a rise in yearly profits. For in the period of total mobilization, when maximizing of essential production was imperative, the Nazis acknowledged by their actions that the traditional or special profit stimuli had proved to be a stronger inducement to managerial efficiency than all Nazi violence and ideals.[8]

How did recognition of the profit stimulus affect price control? When the price commissioner permitted a rise in some prices, many producers clamored for the return to "the free price mechanism." Even the *Frankfurter Zeitung* campaigned for a general rise in food prices in the last days of its existence. But the price commissioner rejected firmly the demand for a general price rise, and insisted that higher prices would merely transfer the excess purchasing power from the buyers to the sellers. The transfer itself would not increase production; it would certainly produce a general price inflation. The subsequent actions of the commissioner and other agencies were able to head off a general price inflation.

The Nazis had to face a peculiar form of inflation. For many Nazis and their economists, inflation was exclusively a price problem. Regulate all markets, fix all prices, then — they believed — inflation cannot occur. Consequently, they decided to leave the traditional income structure undisturbed. The formation of excess purchasing power and monetized capital can be avoided through government borrowing. The Nazis ridiculed all proposals for compulsory savings, and deliberately decided to leave the act of saving uncontrolled. The demand for a special tax upon saving accounts was rejected. Instead, many inducements were offered to increase savings, like the "iron savings" plan and the invitation to deposit depreciation accounts with the government. These schemes offered reduced tax obligations for increased savings, but their success was very limited.[9] The schemes failed to absorb all

7. The squeeze upon the trader's margins, however, was greatly intensified during the last two years of the regime. Cf. "Kostenkontrolle im Grosshandel," Frankfurter Zeitung, September, 12, 1942.

8. The same treatment was refused to laborers and small but nonessential firms; they felt the full force of the increasing Nazi violence. Cf. H. Grimpe, "Der unentbehrliche Kleinbetrieb," Die Deutsche Volkswirtschaft, 1944, p. 696.

9. Otto Nathan, The Nazi Economic System, Durham, 1945, pp. 297 ff.

excess purchasing power, nor did they divert all monetized capital to the treasury. Hence inflation developed in spite of the extensive regulation of the capital and money markets.

The quantity of money in circulation increased greatly. Excessive borrowing by the government destroyed the confidence of many in the stability of the banking system. In spite of many official threats, people began to hoard notes.[1] Many firms invested their funds in liquid assets. This prevented the banks from absorbing all government bonds. The treasury was compelled to incur a rapidly rising short-term debt. Hence distrust in the value of money and government bonds and a change in war finance was caused by the "purely monetary" inflation. Excess purchasing power and liquid assets gradually undermined the monetary system, which began to disintegrate even before the military defeat of the Nazis.

At first the governmental price fixing neutralized the price structure against inflationary pressure, but the uncontrolled funds caused a special price inflation in securities from 1940 to 1942. After some hesitation, the stock markets were extensively regulated like all other markets. Speculation in real estate led to a general prohibition of changes in ownership, except when agriculturally beneficial. In spite of the heavy penalties, barter trade gained in significance. The inflationary pressures proved stronger than the market and price controls of the Nazis. Barter trade established a new market that was beyond the official controls. In 1944 the Nazis realized that they could not suppress barter trade and recognized some forms of such trade, but they fixed the values of goods exchanged. Premium payments were increasingly paid in kind: whoever delivered more milk could receive higher feed rations. Barter exchanges were established for equipment, tools, and also for consumer goods.[2] But control of barter trade was unsuccessful; and barter proved that the controlled price structure could not be immunized against inflation. When the Nazi power crumbled, barter trade superseded regulated marketing, and the controlled price structure disintegrated.

1. See the speech of Funk on "Sparen im Kriege," *Frankfurter Zeitung*, June 29, 1943.

2. "Neue Formen des Warentausches," *Berliner Boersen-Zeitung*, August 13, 1944.

CONCLUSIONS

Under conditions of central economic planning, determination of profits became the function of government. Similarly, the Nazis determined the uses to which these profits were put. In both respects, however, the Nazis failed partially. A ceiling on yearly profits failed because the Nazis were compelled to retain the profit stimulus as an inducement to owners for high efficiency and maximum output of essential goods. Effective control over investment or consumption of capital failed in part because the act of saving was not controlled. "Free" savings and lack of effective control of disinvestment of capital caused excessive purchasing power and monetized capital, in spite of extensive control of money and capital markets, in spite of price fixing and governmental allocation of agents of production. In consequence, the Nazis could not achieve full coördination of profits with central economic planning.

The lack of coördination involved two drawbacks — for economic planning. On the one hand, the volume of property income (e.g. yearly profits) could not be known and determined in advance. This was responsible for the "purely monetary inflation." Actually, it caused the price boom in securities and contributed to the loss of confidence in the banking system. On the other hand, the opportunity for rising yearly profits contributed to a governmentally permitted rise in prices. Hence price and money controls suffered from the lack of coördination between profits and economic planning.

One has to realize, however, that these setbacks were just temporary defects which could have been overcome in the absence of military defeat. Improvements in excess profit taxation could have reduced the dangers springing from the uncontrolled portion of yearly profits. Extension of multiple prices to civilian goods could have reduced the impact of uncontrolled profits upon the price structure. Continued property transfer in conquered lands to nazified concerns would have absorbed much of the monetized capital. Improved methods of borrowing and some new forms of compulsory saving would have neutralized much of the excessive purchasing power, at least as long as the rations of 1940–42 prevailed. None of these improvements could be attained in the period of defeat.

Two inferences can safely be drawn from the Nazi planning

experience. Central planning and private profits developed as the two typical features of modern state capitalism.[3] There is no reason to expect that one will absorb the other, for in the framework of such an economy they can be made compatible; but there will always be practical difficulties in adjusting the recognized profit stimulus and total profits to the requirements of central planning. Such difficulties can be overcome without endangering central planning as long as the ruling party or army is in command of the state and not threatened by defeat.

ARTHUR SCHWEITZER.

UNIVERSITY OF WYOMING

3. State socialism does not face the same problem. In dictatorial state socialism, private ownership in the means of production is eliminated and with it the profit system. Adequate incentives for workers and managers is the problem that remains to be solved. Democratic state socialism retains private ownership, even in nationalized industries, but the owners are reduced to rentiers. The connection of their income with the volume of production and efficiency of management has been severed. Hence in neither system is there, in principle, a need to adjust profits to the requirements of planning.

WAGES, PROFIT AND PRICES

SUMMARY

I. Purpose of the article, 26. — Theoretical background of the analysis, 26. — Disregarding any particular link in the chain of cost-price relationships, 29. — II. Basis of the numerical measures of interdependence used, 30. — Setting up the basic equations, 31. — Results of the analysis, 31. — III. The question of fixed technical coefficients, 38.

The purpose of this article is to present the measures of certain fundamental interrelationships between the wage rates paid, profits derived, and the prices received by all the various branches of the American economy during the last normal prewar year, 1939.

The exposition of the actual numerical result of the study is preceded by a short presentation of the simple — some readers may find it all too simple — theoretical background of the analysis. General comments indicating the principal limitations of the theoretical scheme selected and pointing out the probable nature of bias thus introduced in the numerical findings follow at the end.

The cost-price relationship within a separate industry as conditioned by its technical structure constitutes the factual basis of the whole analysis. Let $P_1, P_2, P_3, \ldots, P_i, \ldots, P_m$ be the prices of the products of the m separate industries, and $a_{i1}, a_{i2}, a_{i3}, \ldots a_{im}$, the technical input coefficients showing the amounts of product of industry 1, industry 2, and so on used in production of commodity i, i.e., the physical amounts of these different kinds of goods absorbed by industry i per unit of its own output. The output of an industry, as referred to in all the following discussion, is defined with exclusion of the products consumed by the same industry in which they have been produced. Thus

$a_{11} = a_{22} = \cdots = a_{ii} = \cdots = a_{mm} = 0$ by definition.

The quantitative relationship between the price P_i of the product of any industry i and the prices of the products of all the other $m-1$ industries used by it can be described as follows:

(1a) $\quad a_{i1}P_1 + a_{i2}P_2 + a_{i3}P_3 + \ldots + a_{im}P_m + R_i = P_i \quad (i = 1, 2, 3, \ldots m)$

This equation also represents a definition of R_i, which is the difference between the price of the commodity produced by the particular industry and that part of its unit costs which consists of payments for materials and services, in short, for all kinds of products purchased from the other industries. In other words, R_i can be thought of as the "value added" originated in industry i per unit of its product. An accountant would recognize in the above equations the familiar unit cost and profit computations used to determine the appropriate price for, say, a next year's model of a Plymouth four-door sedan. One equation of this type can be set up for each of the m industries comprised in our economic system. Considering, now, the technical coefficients — all the small a's — as known, we arrive at a system of m linear equations with $2m$ yet unknown variables: the m prices, P_1, P_2, $\ldots P_m$ and the m values added, R_1, R_2, R_3, $\ldots R_m$.

After having assigned some definite numerical magnitudes to any m of these $2m$ variables, we can solve the system (1a) for the remaining m variables. A price-fixing authority, for example, could calculate the combined amount of wages and profits which each of the m separate industries would derive per unit of its output on the basis of any particular set of prescribed m prices. On the other hand, having first decided on the amounts of combined wage and profit income which each of the m industries should be able to earn per unit of its output, the same authority could, on the basis of given technical relationships, determine the one and only system of prices which would have to be established in order to make the attainment of this particular income allocation possible. As a matter of fact, having arbitrarily fixed any k prices and m-k values added, it is possible to solve our $2m$ equations for the remaining k-m prices and k values added.

The actual solution of system (1a) for all the m prices in terms of the similar number of all the values added can be written down in the form of m equations — one for each of the unknown prices:

(2a) $\qquad P_i = A_{1i}R_1 + A_{2i}R_2 + \ldots + A_{ii}R_i + \ldots + A_{mi}R_m$
$$(i = 1, 2, 3, \ldots m)$$

Each one of the capital A's depends in its magnitude on all the

small a's, i.e., it is a function of the technical structure of all the m industries.[1]

A_{ki} shows the relationship between the price P_i and the value added, R_k, derived by industry k per unit of its output. Should R_k, for example, be increased by one dollar, that is, should the combined wages and profits earned in industry k per unit of its product be increased by one dollar (while the R's, that is, wages and profits, in all other industries remain the same as before) the price P_i of the product of industry i will go up by A_{ki} dollars.[2]

To introduce into our system the wage rate, i.e., the price of labor P_n explicitly, it is only necessary to split the value added R_i of each industry into its two component parts, which in this instance are the wages W_i and the non-wage income or, for short, profits π_i. The wage costs per unit of output can, furthermore, be represented as the wage rate P_n multiplied by the labor input coefficients a_{in}; this coefficient, in analogy with the other technical coefficients, represents the amount of direct labor hired and used by industry i per unit of its output. Thus

1. If D represents the determinant

$$
\begin{vmatrix}
1 & -a_{21} & -a_{31} & .. & -a_{m1} & -a_{n1} \\
-a_{12} & 1 & -a_{32} & .. & -a_{m2} & -a_{n2} \\
-a_{13} & -a_{23} & 1 & .. & -a_{m3} & -a_{n3} \\
. & . & . & .. & . & . \\
. & . & . & .. & . & . \\
-a_{1m} & -a_{2m} & -a_{3m} & .. & 1 & -a_{nm} \\
-a_{1n} & -a_{2n} & -a_{3n} & .. & -a_{mn} & 1
\end{vmatrix}
$$

while D_{nn} stands for the algebraic complement (minor) of the element $-a_{nn}$. and $D_{nn \cdot ki}$ for the complement of the two elements $-a_{nn}$ and $-a_{ki}$, then

$$A_{ki} = \frac{D_{nn \cdot ki}}{D_{nn}}$$

2. Comparing the meaning of A_{ki} as a measure of dependence between the price of commodity i and the income derived by industry k with its significance as a coefficient showing the effect of a change in the "final demand" for commodity k upon the physical output of industry i, we observe the notable parallelism existing between the price variation and physical changes as seen from the point of view of the economic system as a whole. The latter aspect of the problem has been discussed in two articles published in this JOURNAL, February, 1944 and February, 1946.

(3a) $\quad R_i = W_i + \pi_i = a_{in}P_n + \pi_i$ $\qquad\qquad (i = 1, 2, 3, \cdots m)$

Substituting each of these m relationships in the corresponding equations of system (2a) we arrive at a new set of equations:

(4a) $\quad P_i = (A_{1i}a_{1n} + A_{2i}a_{2n} + \cdots + A_{mi}a_{mn})P_n +$
$\qquad\qquad A_{1i}\pi_1 + A_{2i}\pi_2 + \cdots + A_{mi}\pi_m \qquad (i = 1, 2, 3, \cdots m)$

Each of these equations describes the price of one commodity as depending upon the wage rate P_n paid in all industries and the profit rates $\pi_1, \pi_2, \cdots, \pi_m$ earned by each one of them. The expression in parenthesis shows by how much the particular price P_i would go up (or down) for every dollar added to (or subtracted from) the wage rate P_n, this on the assumption that the profits earned per unit of output in all the industries remain the same as before.

These and other similar formulae enable us to determine the total price effect of all possible combinations of separate and simultaneous changes in wage and profit rates of various industries.

If for some reason it were desirable to disregard in our analysis one particular link in the chain of cost-price relationships, this can be done by solving system (1a) for a different set of independent variables and omitting one of the equations. For example, if it is assumed — as it might very well be, from a short-run point of view — that a change in the replacement costs of industrial machinery would be "absorbed" by the machinery-using industries, rather than "passed forward" by means of corresponding adjustment in the prices of finished products, this assumption can be taken care of by shifting the price of machinery, let it be P_2, from the set of dependent into that of the independent variables. The group of m independent variables will consist now of the m-1 other prices and one of the values added, say the value added in the machinery industry, R_2. The price of machinery occupies now, in the solutions of our equations, a position similar to that of the wage rate, P_n, in (4a). If by any chance — and this is frequently the real reason for wanting to disregard these particular elements of the system — the input coefficients describing the amounts of equipment required by various industries for purposes of replacement are unknown, a relationship of the type described by system (2a) can still be established:

(2b) $\quad P_i = \bar{A}_{1i}\bar{R}_1 + \bar{A}_{3i}\bar{R}_3 + \cdots + \bar{A}_{ii}\bar{R}_i + \cdots + \bar{A}_{mi}\bar{R}_m$
$\qquad\qquad\qquad\qquad\qquad (i = 1, 3, 4, \cdots, m)$

Here the replacement expenditures on machinery $a_{12}P_1$, $a_{32}P_3$, $\cdots$, $a_{m2}P_m$ are considered to be a part of the redefined values added $\overline{R}_1$, $\overline{R}_3$, $\cdots$, $\overline{R}_m$ of the respective industries:

(3b) $\overline{R}_i = a_{i2}P_2 + a_{in}P_n + \pi_i$ $(i = 1, 3, 4, \cdots, m)$

The price of machinery, as such, P_2 is eliminated from the set of variables; so is R_2, the value added in the machinery industry. At the same time, the equation describing the cost-price relationship within the machinery industry — the equation which would have contained R_2 — is left out of consideration. Thus system (3b) includes m-1 equations and $2m$-2 variables, of which m-1 are prices and m-1 the redefined values added.

The new short-circuited theoretical description of our economic system takes, as compared with the complete description (2a), a greater number of elements for granted. Thus it is able to explain a smaller number of dependent variables At the same time, it offers the practical advantage of requiring for purposes of actual computations a smaller amount of empirical information than the more comprehensive system (2a).

Once a set of price changes has been computed, it can also be expressed in terms of an index number. Let, for example, ΔP_1, ΔP_2, $\cdots$, ΔP_m be a particular set of price changes computed on the assumption of a general 10 per cent wage rise. If $P_1{}^o$, $P_2{}^o$, $\cdots$, $P_m{}^o$ represent the prices of the base period and a_{n1}, a_{n2}, $\cdots$ a_{nm} the fractions of the average consumer's dollar spent during this base period on commodity 1, commodity 2, and so on — then the weighted average

(5) $\dfrac{\Delta P_1}{P_1{}^o} a_{n1} + \dfrac{\Delta P_2}{P_2{}^o} a_{n2} + \cdots + \dfrac{\Delta P_m}{P_m{}^o} a_{nm}$

will give the price index describing the corresponding change in the average cost of living.

II

The numerical measures of quantitative interdependence between prices, wages and profits shown in the tables and graphs below are computed on the bases of technical relationships characteristic of the American economic system of the last normal prewar year, 1939. The technical input coefficients used in these calculations are derived from a large input-output chart showing the flow of commodities and services between all branches of produc-

tion, transportation, distribution and consumption. The chart itself has been reproduced in a previous article.[3]

In setting up the basic equations all the expenditures on durable equipment, all purchases for purposes of inventory accumulation, as well as tax payments, were treated as part of the values added of the respective industries, that is the *indirect* price effects operating through induced changes in these particular cost items have been neglected. For more or less accidental reasons of computational convenience, payments for imported commodities and services have been treated in a similar way. The loose connection which exists between the import and export prices could be used to rationalize this procedure; furthermore, foreign trade plays a relatively minor part in the American economy, so that a different theoretical treatment of export and import prices would introduce a hardly noticeable change in the numerical results of the following computation.

Table I shows in Column 1 the value of output of the separate goods and services. As stated above, the products sold and consumed within the same industry are excluded from these totals. The other columns show the composition of values added in various industries, as used in our computations.

The final results of the quantitative analysis are presented in Table II. Column 1 shows by how many per cent the price of each kind of goods and services would have increased (or fallen) if the wage rates *in all industries* were raised (or reduced) by 10 per cent — while non-wage incomes as well as all the other components of the values added in all the industries had remained the same as before. Each figure entered in the next column shows the (smaller) price changes which would result, were the wage rates increased by 10 per cent *only in the industry immediately engaged in producing* the particular commodity; the wage in all other industries remaining the same as before. These figures show, in other words, what might be called the *direct* effects of the wage rise in one particular industry on the price of its own output. The

3. "Exports, Imports, Domestic Output and Employment," this JOURNAL, February, 1946. A detailed explanation of the individual classifications used in the construction of all our tables is contained in the appendix to that article. The following minor changes in this classification have been made, however, for the purposes of the present computation: munitions are included in Chemicals and not in Metal Fabricating; Restaurants are taken from Unallocated, and classed with Trade.

TABLE I

OUTPUT PRODUCED, INCOME DERIVED AND SELECTED EXPENDITURE ITEMS OF AMERICAN INDUSTRIES FOR THE YEAR 1939

Unit: One Million Dollars

Industry	Total Value of Product 1	Wages and Salaries 2	Non-Wage Incomes 3	Total Income = Cols. 2+3 4	Expenditures on Capital Goods 5	Addition to Stocks and Inventories 6	Taxes 7	Imports 8	Total "Value Added" =Cols. 4+5+6+7+8 9
1. Agriculture and Fishing	10,121	1,102	3,927	5,029	954	0	512	337	6,832
2. Food Processing	13,282	1,526	653	2,179	154	24	1,583	824	4,764
3. Ferrous Metals	2,622	756	127	883	71	29	108	22	1,113
4. Motor Vehicles, Industrial and Heating Equipment	4,975	1,575	553	2,128	126	7	302	10	2,573
5. Metal Fabricating	6,970	2,395	439	2,834	139	36	350	18	3,377
6. Nonferrous Metals and Products	1,616	418	198	616	19	48	95	331	1,109
7. Stone, Clay and Glass Products	2,083	563	167	730	34	17	74	63	918
8. Fuel and Power	10,988	2,596	1,175	3,771	1,058	28	1,766	72	6,693
9. Chemicals (and Munitions)	3,434	647	401	1,048	44	33	185	161	1,471
10. Lumber, Paper, Printing, Publishing	6,411	2,218	396	2,614	140	13	233	259	3,259
11. Textiles and Leather	7,631	2,329	318	2,647	134	33	177	381	3,372
12. Rubber	892	229	53	282	18	0	59	196	555
13. All Other Manufacturing	1,671	633	153	786	14	2	65	61	928
14. Construction	10,089	3,783	353	4,136	268	0	90	0	4,494
15. Transportation	7,477	3,398	1,219	4,617	1,102	0	854	284	6,857
16. Trade and Restaurants	20,723	7,239	3,153	10,392	372	0	1,014	0	11,778
18. Business and Consumer Services	18,525	3,948	6,668	10,616	1,148	0	2,122	0	13,886
21. All Other Goods and Services	22,192	1,579	871	2,450	2,026	0	1,166	92	5,734

TABLE II

PRICE CHANGES RESULTING FROM ASSUMED 10 PER CENT INCREASES IN WAGE RATES AND NON-WAGE INCOMES, COMPUTED ON THE BASIS OF STRUCTURAL RELATIONSHIPS PREVAILING IN THE AMERICAN ECONOMY OF THE YEAR 1939

Unit: One Per Cent

Commodities and Services Produced By:	Price Change Resulting from a 10% Wage Rise in All Industries 1	Price Change Resulting from a 10% Wage Rise in the Industry Producing the Commodity 2	Price Change Resulting from a 10% Rise in Non-Wage Income in All Industries 3	Price Change Resulting from a 10% Rise in the Non-Wage Income in the Industry Producing the Commodity 4	Price Change Resulting from a 10% Rise in Agricultural Wage and Non-Wage Income 5
1. Agriculture and Fishing	2.55	1.12	4.74	3.98	5.10
2. Food Processing	3.23	1.18	2.80	0.51	-1.88
3. Ferrous Metals	4.98	2.92	1.45	0.49	.04
4. Motor Vehicles, Industrial and Heating Equipment	5.56	3.21	2.06	1.13	.05
5. Metal Fabricating	5.54	3.53	1.56	0.65	.05
6. Nonferrous Metals and Products	3.87	2.60	1.81	1.23	.03
7. Stone, Clay, Glass Products	4.98	2.72	1.84	0.81	.05
8. Fuel and Power	4.28	2.47	1.88	1.12	.03
9. Chemicals (and Munitions)	4.14	1.92	2.50	1.19	.31
10. Lumber, Paper, Printing, Publishing	5.20	3.53	1.56	0.63	.22
11. Textiles and Leather	4.91	3.08	1.72	0.42	.56
12. Rubber	4.21	2.58	1.45	0.60	.08
13. All Other Manufacturing	5.41	3.81	1.75	0.92	.09
14. Construction	6.44	3.75	1.41	0.35	.15
15. Transportation	5.21	4.67	1.92	1.68	.01
16. Trade and Restaurants	4.95	3.61	2.50	1.57	.24
18. Business and Consumer Services	3.23	2.19	4.08	3.70	.04
21. All Other Goods and Services	3.19	0.82	1.66	0.45	.13

differences between these figures and the corresponding entries in Column 1 measure the *indirect* effects, on the price of the products of the industry concerned, of a 10 per cent wage rise in all the other industries. The indirect effects are those transmitted through increased costs of materials and supplies purchased from other industries.

Columns 3 and 4 in Table II show the total and the direct effects of a hypothetical 10 per cent increase in profits, interest and other types of non-wage income, computed on the assumption that wages as well as the other four components of the values added in all the industries remain constant. The differences between these two parallel sets of figures obviously measure the indirect dependence of the unit cost of production of each industry upon profits and other non-wage incomes included in the prices of output of all the other industries. Figures 1 and 2 facilitate comparison of the different kinds of price effects presented in Table II.

Lack of space makes it impracticable to tabulate the magnitudes of all the partial relationships between the price of every commodity, on the one hand, and the wage and non-wage incomes earned in each one of many industries, on the other. As a practically important example of this type of dependence, the computed price effects of a 10 per cent increase in total (wage and non-wage) income derived by Agriculture per unit of its output are listed in Column 5 of Table II, and also represented graphically in Figure 3. In this connection it should be observed that the distinction between the wage and the non-wage income in Agriculture does not coincide with the division between labor and non-labor income as it does in most other industries.

The last right-hand bar in each of the three graphs represents the average of the separate price changes shown in the same diagram. With all separate price changes weighted in proportion to the fractions of the consumer's dollar spent in the year 1939 on each kind of commodity, one of these averages shows that a 10 per cent increase in the wages paid by all industries per unit of their respective products would raise the cost of living by 3.7 per cent, the other indicates that a proportionally similar increase in non-wage earnings would reduce the purchasing power of the consumer's dollar by 2.6 per cent, and the last one means that a 10 per cent

Fig. I. Price Changes Resulting from an Assumed 10 Per Cent Rise in Wage Rates Computed on the Basis of Structural Relationships Prevailing in the American Economy of the Year 1939
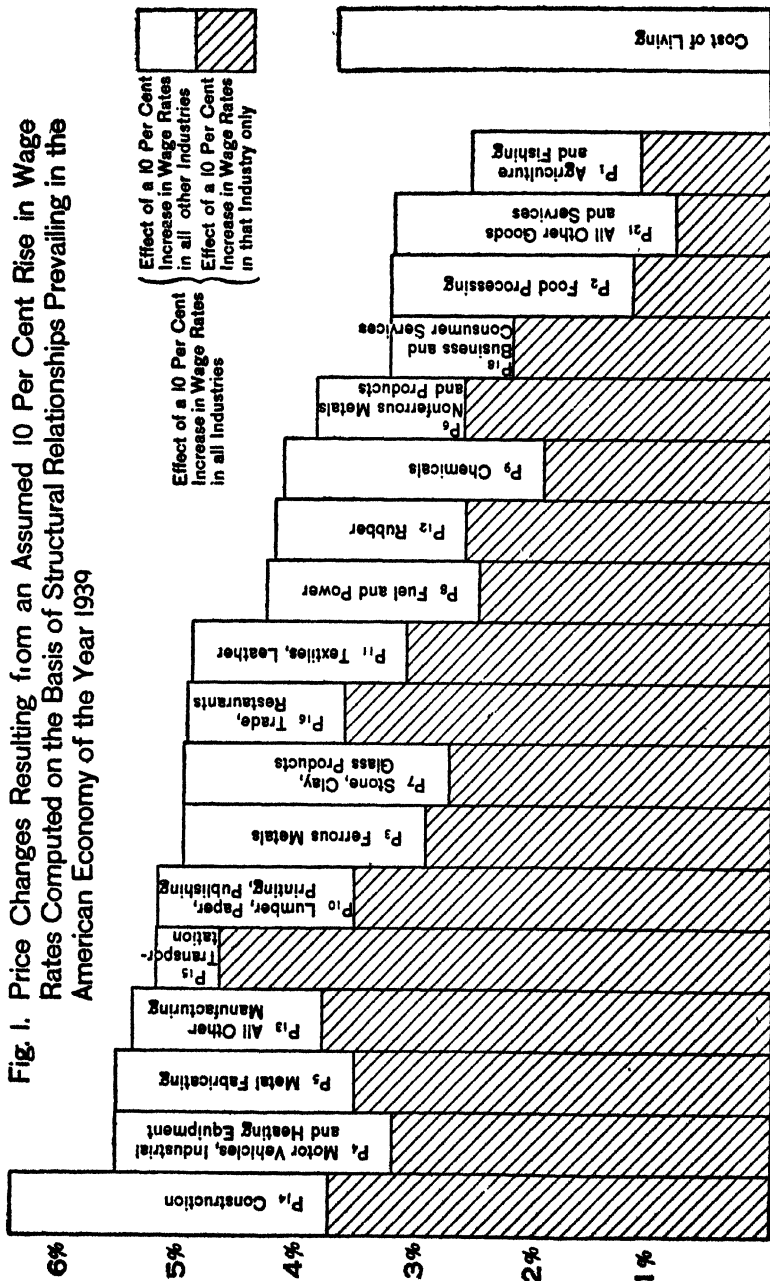
Fig. 2 Price Changes Resulting from an Assumed 10 Per Cent Rise in Profits and Other Non-Wage Incomes Computed on the Basis of Structural Relationships Prevailing in the American Economy of the Year 1939

Fig. 3. Price Changes Resulting from an Assumed 10 Per Cent Rise in Agricultural Wage and Non-Wage Income Computed on the Basis of Structural Relationships Prevailing in the American Economy of the Year 1939
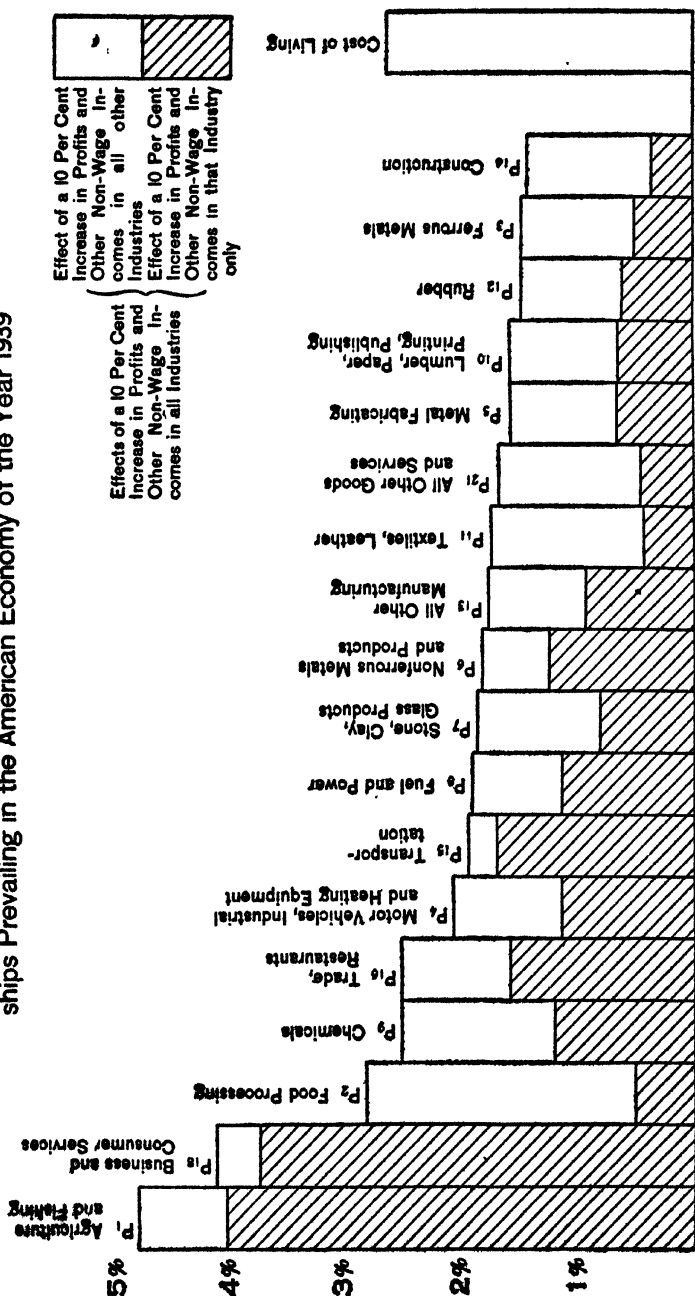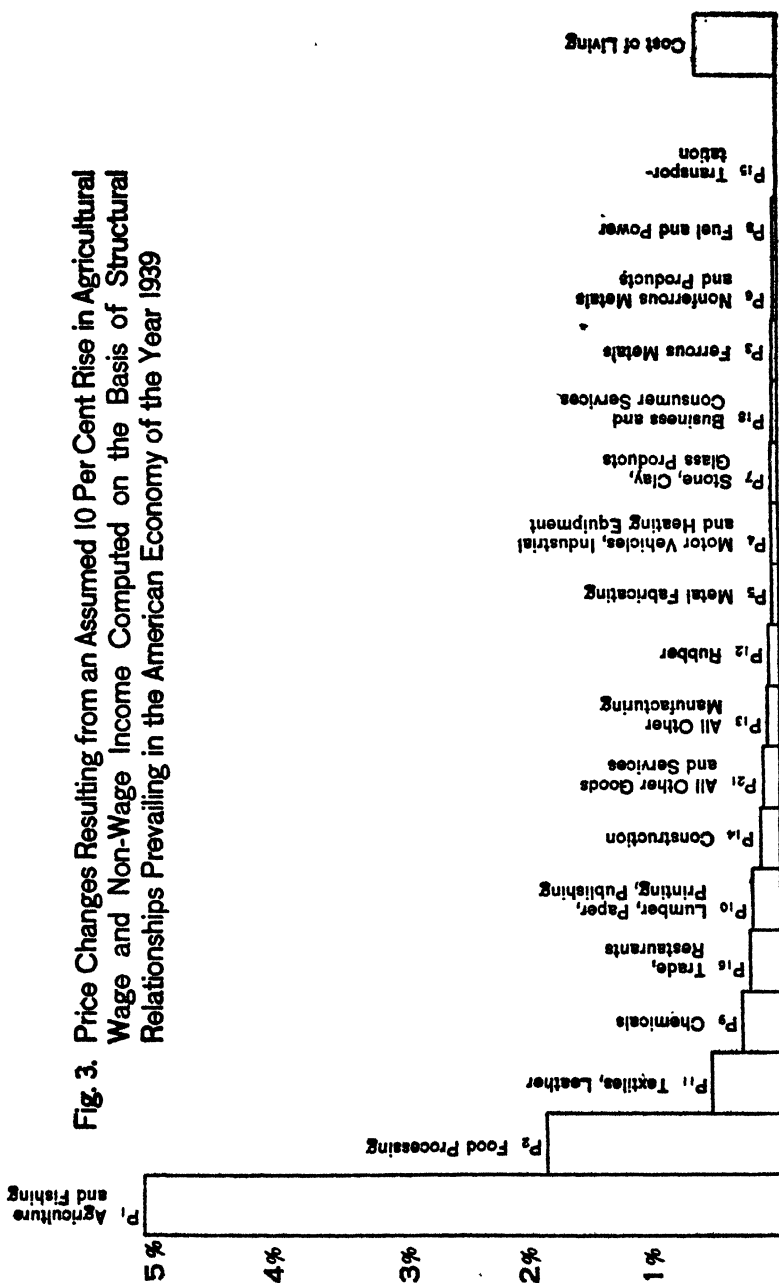
increase in all agricultural incomes derived per unit of farm products would have added to the cost of living 0.7 per cent.

Since, as has been pointed out above, the distinction between wage and non-wage income in Agriculture has a very special meaning, the first two averages have also been computed with exclusion of products of Agriculture and of the Food Processing industries. The non-agricultural component of the cost-of-living index (taken separately) indicates a 3.8 per cent rise in consequence of a 10 per cent wage increase, and 2.4 per cent are added to it as a result of a general 10 per cent increase in all non-wage incomes. As should be expected, the exclusion of agricultural prices makes the cost-of-living index more sensitive to wage changes and ess dependent upon variations in non-wage incomes.[4]

If wage and non-wage incomes constituted the only components of all the values added, as used in these price computations, a simultaneous and proportionally equal increase in both types at earnings would naturally raise, under our assumption, the costs of living in the same ratio. The assumed constancy of the depreciation costs, of the tax burden, as well as all the other non-income elements of the values added, is responsible for the fact that the assumed increases of wage and of non-wage incomes add in their combined effect 6.3, rather than 10 per cent, to the average costs of living.

III

The assumption of fixed technical coefficients, which constitutes the basis of the foregoing empirical analysis, can be questioned from the point of view of general theory of production. Insofar as the proportions in which the separate factors can be

---

4. In computation of the changes in the costs of living, all consumers' expenses, including those which like the tax burden are assumed to be constant per unit of output, are included in the respective averages. If only those consumers' expenses were taken into consideration which are allowed according to our assumptions to change, the computed increases in the cost of living would obviously be larger: 4.1 per cent for the result of a 10 per cent increase in wages and 2.9 per cent for the corresponding increase in the non-wage income. A 10 per cent rise in agricultural income would increase the variable part of the cost of living by 0.72 per cent, which can be compared with the 0.65 per cent increase in the total index (in the text the latter figure has been rounded out to 0.7 per cent). Applying the same argument to the non-agricultural component of the cost-of-living index, i.e. excluding from it the tax payments and other such items, we raise the 3.8 per cent to 4.4 per cent and the 2.4 per cent to 2.8 per cent.

combined within the same production function (i.e., at any given state of engineering information) are variable, these proportions will most probably vary with every change in their relative prices. This theoretical proposition so clearly stated by Pareto in his criticism of Walrasian fixed coefficients of production is beyond dispute. It is, however, not the fundamental validity of the principle of substitution but its quantitative significance which is important from the point of view of empirical analysis. The smaller the variation in production coefficients induced by any given range of changes in factor prices, the smaller obviously will be the empirical error introduced in our computations by the assumption of invariable input ratios. Lack of detailed statistical data covering sufficiently wide sectors of the national economy makes it impossible to determine the actual magnitude of errors introduced in our quantitative findings. An indirect test presented elsewhere[5] seems to indicate that, within the range of price changes which actually have taken place between 1929 and 1939, these errors lie within relatively narrow limits. So far as the direction of these errors is concerned, on general theoretical grounds one can state that neglecting the principle of substitution we arrived at computed price changes somewhat in excess of those which actually would have resulted from the assumed income variations.

For purposes of prognostication the possibility of large spontaneous changes in the magnitude of technical input coefficients would, of course, be as damaging as any secondary change of comparable magnitude induced through price variations. Here again the indirect test referred to above seems to justify expectation of a considerable degree of stability in the general pattern of price relationships. The results of a more detailed statistical investigation into the nature of technological changes since 1919 will be presently available. Without conclusive factual evidence, the reader's judgment on this controversial problem is certainly as good as the author's, and for obvious reasons it might even be less prejudiced.

WASSILY LEONTIEF.

HARVARD UNIVERSITY

5. See the first of the two articles referred to in footnote 2, p. 28.

SAVINGS IN A STATE
WITH A STATIONARY POPULATION[1]

SUMMARY

I. Introduction: the problem, 40; France a good test case, 41; use of terms, 42. — II. Methods used in computing growth of savings, 43. — Comparative aspects, 43. — III. Growth of savings: France, 44; Germany, 46; Great Britain, 48. — IV. The French reputation for thrift, 50. — Reasons for lower savings in France: distribution of income, 51; reinvestment in business, 54; age distribution, 55. — V. Savings and foreign investments: France, 58; Great Britain, 60; Germany, 61. — Conclusions, 61. — Appendix, 63.

I. INTRODUCTION

During the period from 1871 to 1911, a period of relative normalcy in Western Europe, France saved on the average around 10 per cent of her national income, while England and Germany saved, respectively, 12–15 and 15–20 per cent of their national incomes. It is also estimated that during the same period France placed approximately half of her savings abroad, England about one-fourth, and Germany only around 10 per cent. These phenomena were accompanied by various population trends in the three countries. During the four decades under discussion, Germany and England experienced rapid substantial growth of population, while France remained almost stationary (Table I).

Since all Western countries have begun in recent decades to experience a trend towards a declining rate of population growth, this problem has become a topic for serious discussion on the part of leading economists. In his General Theory and elsewhere, Keynes points out that the cessation of population growth will close one of the main sources of the extensive demand for capital, and as a result the "marginal efficiency" of capital will fall below the level of interest rates. This, in its turn, will entail considerable unemployment. Professor Hansen is also concerned with the diminution of investment openings as a result of the decline in the growth of population. He suggests that "with a stationary population we could maintain as rapid a rise in per capita real income as that experienced in the past by making annually only half the

1. I am deeply indebted to Professor A. H. Hansen for many valuable suggestions and comments.

TABLE I

POPULATION IN FRANCE, GERMANY, GREAT BRITAIN (*a*)

(thousands)

	1870	1890	Per Cent of Growth Between 1870–90	1910	Per Cent of Growth Between 1890–1910	Per Cent of Growth Between 1870–1910
France.......	36,103(*b*)	38,343	(6.2)	39,602(*c*)	(3.3)	(9.4)
Germany.....	41,059	49,428	(20.4)	64,926(*d*)	(31.4)	(58.1)
Great Britain .	26,288	33,253	(26.5)	41,309(*d*)	(24.2)	(57.1)

(*a*) Encyclopédie Française, VII, 7180–7182.
(*b*) French Census of 1871.
(*c*) This increase of 3½ million between the years 1871 and 1910 is to an important extent due to the net immigration during this period.
(*d*) While the percentage growth of population between 1870 and 1910 was about the same in both England and Germany (57 and 58, respectively), there had been a consistently higher birth rate in Germany than in England. The latter had achieved her substantial increase in part as a result of a lower mortality rate. Since the concept of a dynamically growing population is most often identified with a higher birth rate, Germany was in this respect more favorably situated than England (see Appendix for birth and mortality rate tables). Furthermore, Germany achieved a considerably higher proportionate growth during the last two decades under study, 31.4 per cent increase as compared with 24.2 per cent for England.

volume of new investment to which we have been accustomed."[2]

It is the purpose of this article to test empirically the impact of relative stationariness of population upon savings in a large Western country in a state of "normalcy." France during the years 1871–1911 seems particularly well suited for such a test. For comparison this analysis is accompanied by a similar study of Germany and the United Kingdom for the same period. The two countries were selected because of the striking similarity between the economic structures and the level of material civilization of all three during the years under study.

The years 1871–1911 were chosen because the French population remained almost stationary during the four decades, and France underwent no change in territorial status. (The important industrial provinces of Alsace-Lorraine were lost in 1870 and regained in 1919.) The period was also characterized by an absence of major perturbations, such as wars or revolutions in Western Europe. Economic life in all the three countries flowed smoothly; the improvement in the standard of living proceeded gradually. The relative absence of trade barriers, freedom of migration of both men and capital, accessibility of scientific information — all

2. A. H. Hansen, *Readings in Business Cycle Theory*, Philadelphia, 1944, p. 381.

of these and other economic and political factors provided an excellent atmosphere for the economic development of each of the three nations.

As indicated above, Germany, which experienced the largest expansion in population in the years 1871–1911, especially between 1890–1911, also consistently saved a higher percentage of her annual income than either France or England. The proportion of national output saved appears to be closely correlated with the population trends, with the result that stationary France saved the lowest share of national income.

This study also lends support to the thesis advanced by Keynes and Hansen that a stationary or declining population faces the problem of an excess supply of capital, since stationary France of 1871–1911, in spite of saving a smaller share of her national income than Germany, was confronted with a relative abundance of capital, which could not find profitable employment within the country.[3]

A phenomenon of major importance which was different in the case of one of the three countries, as opposed to the other two, relates to population trends. In the following pages the relationship between the population movements and savings will be analyzed, as well as the methods and reasoning used in arriving at the above findings.

For the purpose of this study, the terms "savings," "wealth," and "capital" are treated as synonymous. This is done in part because it is the widely accepted procedure to regard national savings as accumulations of wealth and capital,[4] and also because this has been the treatment accorded to the three concepts in the literature dealing with savings and national wealth for the period under consideration. Throughout this study the term "savings" is used in the sense of net savings, since the concept of gross savings is of relatively recent origin and was almost never mentioned in the literature of the period on which this paper is based.

3. This, of course, is not the only explanation for the large exports of French capital. As indicated elsewhere, political considerations also served as inducements for such exports. There is, however, ample ground for the belief that the cessation of population growth curtailed the extensive demand for capital, thus giving impetus to the great exodus of capital from France into foreign investments.

4. Conference on Research in National Income, Vol. II, S. Kuznets, p. 47; Morris Copeland, p. 75; entire volume passim. F. von Hayek, Encyclopedia of the Social Sciences, Vol. XIII, p. 548.

II. Methods Used in Computing the Growth of Savings

There are two basic approaches to the problem of measuring national savings. One method consists of studying the annual accumulation of wealth with the aid of censuses of production, transportation and distribution. The other method consists of comparing national wealth figures at successive points of time. Both procedures have their able defenders. In this study, however, the choice of method is dictated by the fact that the estimates of national savings in each of the respective countries for the years 1871–1911 are based almost entirely on the estimates of wealth during certain years of that period. This method was used by the leading authorities on national income and savings, by Giffen in England, by Pupin in France, and by Helfferich in Germany.

It should also be emphasized that our interest is concentrated, not so much on the absolute amounts of savings accumulated, data which are admittedly inadequate, as on the movement of the savings. The purpose is to ascertain whether and to what extent the population trends in each of the countries studied affected the growth or decline of savings. With this aim in view, the computation of the exact amount of national wealth accumulation at a given period is of still less significance.

The test of comparative growth of savings cannot be the absolute sums saved, since in the case of two countries starting with approximately the same population, other things being equal, the one with a faster growing population would naturally be in possession of more total capital than the country with stable numbers of population. Neither would a real trend of capital accumulation be revealed by a study of per capita savings, since in a country with a higher average income more will be saved per capita than in one having a lower average income. In our judgment, the best criterion of comparative wealth accumulation as between different countries is the percentage of national income saved within each state.

Our study of comparative wealth increase is facilitated by the similarity of the trend of prices in all three countries for the years 1871–1911, as may be seen from Table II. The effects of changes in prices on capital valuation may, then, for our purposes be assumed to have been the same in France, Germany and the United Kingdom.

TABLE II

INDICES OF WHOLESALE PRICES(a)

(1890–1900 Average = 100)

Year	France(b)	Germany(c)	U.K.(d)
1872...............	159	155	164
1876...............	144	134	143
1881...............	130	127	127
1886...............	106	101	104
1891...............	109	113	108
1896...............	91	93	92
1901...............	105	115	106
1906...............	116	112	106
1909...............	116	112	112

(a) Bulletin de la Statistique Générale de la France, Vol. 1, p. 68.
(b) Based on prices of 43 imported commodities
(c) Based on prices of 28 imported commodities at Hamburg until 1888; of 42 commodities since 1889.
(d) Sauerbeck's index numbers for 45 articles.

III. THE GROWTH OF SAVINGS

France. The prevailing method of calculating national capital in France has been that based on the "annuité successorale,"[5] which may be translated as the amount annually transmitted by inheritance within a country. The introduction of this method of estimate in France was prompted by the lack of income tax data, and by the possession of detailed records on inheritances. From the point of view of this particular study, it is worth reiterating that our concern is not so much with the exactness of the absolute amounts as with the consistency of the procedure employed by the sources used in this study, which will permit us to judge the evolution of national wealth over a period of years.

In his investigation of the private wealth of France for the years 1878–1911, René Pupin,[6] a leading French authority on income and savings, employed the method of "annuité successorale" and that of direct evaluation. He arrived at results which are summarized in Table III. According to his figures there was an increase in French savings between 1878 and 1903 of about 45 billion francs. A detailed analysis of the changes in the value of French wealth for that period led Pupin to raise the estimate of the increase of savings to 51–52 billions. He concludes that the for-

5. In the English economic literature this procedure is referred to as the devolution-interval method.
6. René Pupin, La Richesse de la France Devant La Guerre, Paris, 1916.

TABLE III

Estimate of French Private Wealth
(billion francs)

Year	Through method of "Annuité Successorale"	Through Direct Evaluation
1878...............	209.2	204
1883...............	220.1	
1893...............	251.4	250
1903...............	254.8	
1911...............	283.4	286.3

mation of new capital during the period 1878–1902 proceeded at an annual average rate of a little over two billion francs a year, an amount equivalent to 8–10 per cent of the average national income for the years 1878–1902.[7] Following a similar method of computation, he concludes that the rise in wealth from 1903 to 1911 was around 29 billion francs, amounting to an average annual savings of 3.6 billion. In view of the larger national income of France during the first decade of the twentieth century, annual French savings still hovered around 10 per cent of the total national income. Pupin thinks that the years between 1885 and 1910, especially 1895 to 1910, were exceptionally favorable for savings in France, but he adds wistfully, "We have not reached the rate of our grandfathers, who saved 13 per cent of their national income."[8]

Other French economists came to a close agreement with Pupin on French national wealth and savings. E. Théry,[9] who devoted several studies to the subject of French wealth, used the method of direct evaluation. Most of his computations deal with the years 1892–1908; his estimates for these years are 243 billion and 287 billion francs, respectively. This represents an average annual growth of savings of around 2.7 billion francs, an amount not far from Pupin's figures for the same period.

Robert Lascaux gives detailed estimates on French wealth for the years 1895–1914.[1] His information is gathered from various sources, including that of the "annuité successorale," and his figures are 229 and 282.8 billion francs for 1895 and 1914, respectively. This is also equivalent to an average annual growth of

7. For figures on national income see Appendix.
8. R. Pupin, op. cit., p. 128.
9. E. Théry, La Fortune Publique de la France, Paris, 1911.
1. R. Lascaux, La Production et La Population, Paris, 1921.

capital of 2.7 billions, the same as that indicated by Théry. It is Lascaux's opinion that around 1895 these savings were in the neighborhood of 2.5 billion francs, and that toward 1914 they rose to three billion. Since Lascaux's estimate of French national income for the same two years is 27.5 and 31 billions, respectively, it appears that he, too, places the ratio of national savings to national income at 9 to 10 per cent, a conclusion similar to that arrived at by Pupin. There is, thus far, agreement among leading authorities on French national income and savings that during the four decades preceding 1914 the French saved, on the average, 10 per cent of their yearly national income. (See Appendix.)

A summary of the estimated approximate proportion of savings to the national income of France is presented in Table IV.

TABLE IV

ESTIMATED ANNUAL SAVINGS OF FRANCE AND THEIR RATIO
TO THE YEARLY NATIONAL INCOME

Years	Authority	Estimated National Income (in billion frs)	Estimated Average Annual National Savings (in billion frs.)	Proportion of Savings to Income
1878–1902	Pupin	21.9–27.8	2.04	9–7.5
Around 1895	Lascaux	27.5	2.50	9
1892–1908	Théry	....	2.7	....
1903–1911	Pupin	27.8–32	3.5	11–13
Around 1914	Lascaux	31	3	10
1893–1913	Moulton and Lewis(a)	....	2	...
Pre-1914	A. Neymarck(b)	....	2 to 2.5 (sometimes 3 a year)	...

(a) H. G. Moulton and C. Lewis, The French Debt Problem, New York, 1925, Appendix A.

(b) A. Neymarck, Journal de la Société de Statistique de Paris, 1906, p. 176.

Germany. The most frequently cited source[2] of information on the national income and wealth of Germany before the first World War is K. Helfferich's Deutschlands Volkswohlstand, 1888–1913, translated into English under the title, Germany's Economic

2. E. Crammond, "The Economic Relations of the British and German Empires," Journal of the Royal Statistical Society, 1914, p. 800; J. Stamp, Studies in Current Problems in Finance and Government, London, 1924,

Progress.[3] This work was written in a spirit glorifying the reign of Wilhelm the Second and the great strides made by the German economy during the 25 years preceding 1914. The book bristles with boastful assertions concerning German economic progress, but it gives evidence of careful analysis and the choice of the more reliable methods and data.

In his computations of the increase of wealth in Germany during the period 1896–1912, Helfferich made use of two methods which served as mutual checks upon one another. One procedure was based upon the assessment of the property tax, for which exact statistics were available. The other method was founded on fire insurance data, which have been very carefully recorded in Germany. Josiah Stamp[4] and Robert Lascaux[5] both give credence to Helfferich's figures. Table V is compiled from Helfferich's valuations.

TABLE V

YEARLY INCREASE IN NATIONAL WEALTH OF GERMANY
(billion marks)

Estimates Based upon Property Assessment Taxes		Estimates Based upon Property Insured Against Fire	
Years	Annual Average (billion marks)	Years	Annual Average (billion marks)
1896–1898.......	4.8	1896–1901.......	5.8
1899–1901.......	3.6	....	..
1902–1904.......	5.0	1902–1904.......	7.1
1905–1907.......	6.9	1905–1907.......	8.2
1908–1911.......	10.0	1908–1911.......	9.8

Helfferich places the automatic increase in the value of land and other existing properties during the years 1908–1911 at one and one-half to two billion marks annually. If we are to deduct the latter amount from the estimated annual increase of 10 billions for the same period, this still leaves an annual national accumulation of eight billion marks.[6] Since Helfferich's estimate of German

p. 299; H. G. Moulton and C. E. McGuire, Germany's Capacity to Pay, New York, 1923, p. 190; Robert Lascaux, op. cit., p. 221.

3. K. Helfferich, Germany's Economic Progress, New York, 1914.

4. J. Stamp, Studies in Current Problems in Finance and Government, London, 1924, pp. 466–467.

5. R. Lascaux, La Production et La Population, Paris, 1921, p. 221.

6. It is worth noting that Stamp considers the estimate of 10 billions a close approximation to truth. Journal of the Royal Statistical Society, 1919, p. 467. This would imply an even higher percentage of national income saved than advanced here.

national income for those years is 40 billions, this implies that the German people during the years preceding World War I saved annually about 20 per cent of their national income. Should we accept the higher estimate of 50 billion, as advanced by the Reichsamt for 1913,[7] the percentage of national income saved would still amount to 16–17 per cent per year, a very high ratio indeed.

Table VI gives the ratios for the other years for which Helfferich has provided estimates of savings and income. These indicate

TABLE VI

ESTIMATED RATIO OF GERMAN NATIONAL SAVINGS TO NATIONAL INCOME

Authority	Year	National Income	Estimated Savings (Helfferich)	Approximate Ratio
Schmoller	1895	25 billion	4.8–5.8 (5.3) billion	20%
Hay & Reichsamt	1900	31 billion	3.6–5.8 (4.7) billion	15%
Reichsamt	1907	37.7 billion	6.9–8.2 (7.5) billion	20%
Helfferich	1911	40 billion	8	20%
Reichsamt	1913	50.1 billion	8	16%

that for the two decades preceding the first World War, Germany, with a rapidly growing population, saved a larger share of her national income than did France with a static population during the same period.

The United Kingdom. The generally accepted method of determining the national wealth in Great Britain has been the capitalization of income tax data. To this total are usually added estimates for other items in order to obtain a complete description of the nation's capital. Table VII gives estimates on national wealth prepared by the London Economist and several leading English authorities. In arriving at the estimates of total savings (column six), we made use of the valuations of national wealth given by Giffen and the London Economist, which provided a set of consistently worked out data lending themselves to definite comparability between the years 1875 and 1908. The other estimates given in column two are those made by leading economists, and serve to confirm the figures advanced by Giffen and the Economist. For the years 1909–1914, the totals of the Economist for 1909 and of L. Chiozza Money for 1914 were used.

7. Das deutsche Volkseinkommen vor und nach dem Kriege, Bearbeitet im Statistischen Reichsamt, Berlin, 1932.

TABLE VII

ᴀᴘᴘʀᴏxɪᴍᴀᴛᴇ Eꜱᴛɪᴍᴀᴛᴇꜱ ᴏꜰ Nᴀᴛɪᴏɴᴀʟ Sᴀᴠɪɴɢꜱ ɪɴ ᴛʜᴇ
Uɴɪᴛᴇᴅ Kɪɴɢᴅᴏᴍ ᴀɴᴅ Tʜᴇɪʀ Rᴀᴛɪᴏꜱ ᴛᴏ Nᴀᴛɪᴏɴᴀʟ Iɴᴄᴏᴍᴇ
(Million Pounds)

National Wealth			Approximate Savings			(7)
(1)	(2)	(3)	(4)	(5)	(6)	Ratio of Savings
Year	Estimate	Authority	Interval Years	Amount Saved	Annual Average	to Income(a)
1875	8,548	Giffen(b)	1875–1884	1,492	149.2	12%
1885	10,037	Giffen	1885–1894	626	62.6	4%
1895	10,663	Economist(c)	....	....	....	....
1896	10,125	Milner(d)	1895–1904	2,373	237.3	13.5%
1902	11,413	Moley(d)	....	....	....	....
1905	12,671	Fabian Society(d)	....	....	....	....
1905	13,036	Economist	1905–1908	950	237.5	12%
1908	13,762	Money(e)	....	....	....	....
1909	13,986	Economist	....	....	....	....
1913	16,472	Grammond(f)	1909–1914	approx. 2,000	333.3	15%
1914	16,000	Money(e)	....	....	....	....

(a) These ratios were derived by dividing Average Annual Savings (column 6) by Estimated Annual National Income for corresponding years (see Appendix).
(b) R. Giffen, The Growth of Capital, London, 1889.
(c) Economist, November 25, 1911, p. 1067.
(d) J. Stamp, British Incomes, p. 406.
(e) L. Ch. Money, Riches and Poverty (London, 1910), p. 65.
(f) Journal of the Royal Statistical Society, 1914, p. 801.

On the basis of this table, we find the total growth of savings for the period 1875–1914 to be 87 per cent, but the per capita increase of wealth for the same period,[8] using the same estimates of total wealth, was 35 per cent, the respective amounts for per capita wealth being approximately 260 pounds in 1875 and 350 pounds in 1913. This means that per capita wealth in the United Kingdom grew at a rate of close to one per cent a year.[9]

It is interesting to compare these findings with a study made by Paul Douglas on the growth of "real capital" in the United

8. The population figures for these estimates were taken from the Statistical Abstracts for the United Kingdom for various years.
9. It is interesting to record that Keynes found that the standard of living in Great Britain had risen during 1860–1913 by 60 per cent — a little over one per cent a year for that period. "Galton Lecture," Eugenics Review. April, 1937, p. 15.

Kingdom.[1] In this study Douglas compared the various estimates of capital in this country for the years 1865 to 1914. He did not include in this concept the value of agricultural land, dwelling houses, furniture and similar movable capital. The results were expressed in terms of the 1865 price level, and the following index of real capital per head was obtained:

| 1865 — 100 | 1885 — 154 | 1905 — 174 |
| 1875 — 135 | 1895 — 163 | 1909 — 178 |

This means that the per capita amount of real capital, in the sense defined by Douglas, grew during the 34 years (1875–1909) by 32 per cent; our table of the per capita growth of wealth in a wider meaning of the term yields a per capita increase of 35 per cent for the 39 years 1875–1914, a remarkably close agreement which lends additional support to the probable accuracy of the estimates of the growth of savings in England.

IV. Saving in France

The preceding sections have revealed that the ratio of savings to national income in France, the United Kingdom and Germany were approximately as summarized in Table VIII. The most striking feature of this table is the relatively lower percentage of savings consistently effected by France as compared with those

TABLE VIII

Ratio of Savings to National Income

France		United Kingdom		Germany*	
Years	Ratio of Savings to Nat. Income	Years	Ratio of Savings to Nat. Income	Years	Ratio of Savings to Nat. Income
1878–1902	9–7.5	1875–84	12	1895	20
1903–11	11–13	1885–94	4	1900	15
1914	10	1895–1904	13.5	1907	20
		1905–08	12	1911	20
		1901–1914	15	1913	16

* Professor Hansen in commenting on an earlier draft of this paper expressed his belief that the larger proportion of income saved in Germany as compared with France may in part be due to the greater expansion of bank credit and of the money supply resulting from the greater expansion and growth experienced in Germany. He also agrees, however, that while savings from disposable income may be less divergent than those given by the author for the two countries, the difference between them would still be very considerable.

1. American Journal of Economic and Business History, August, 1930. Quoted by C. Clark, National Income and Outlay, London, 1937.

achieved by Germany and England. According to these data, during the years 1875–1914, France saved a somewhat smaller share of her national income than England, and a much lower proportion than Germany.

The above finding is the more interesting since France has always been considered the saver-nation par excellence. The literature on population in France since 1850 abounds with references to excess savings, which were frequently ascribed to the low birth rate. Jacques Bertillon,[2] Charles Gide,[3] and numerous other writers all played different variations on the same theme, namely, that French savings were growing at the expense of unborn children.[4] Lescure[5] and Pupin[6] advanced various reasons for the alleged superior growth of French savings. Nor is this reputation undeserved, despite the record since the 1870's. It is due, in our opinion, primarily, to the more equitable distribution of wealth and income in France, which implies that the act of saving was predominantly performed by the masses of the people, rather than by the corporations and the less numerous affluent classes, as in England and Germany.

One possible explanation of the lower savings in France is the more equal distribution of the national income. We have enough evidence to conclude that France suffered less from income and wealth inequality than either Great Britain or Germany. Pierre Leroy-Beaulieu[7] points out that in 1905 the very large incomes, exceeding 100,000 francs a year, amounted in England to 3.5 billions, in Germany to 2.5 billions, and in France to 572 millions. This represented, respectively, 6, 6.5, and 2.5 per cent of the total

2. J. Bertillon, La depopulation de la France, Paris, 1911, passim.

3. Charles Gide, Revue Economique Internationale, March, 1910.

4. An example of how far this attitude can be carried may be gleaned from this statement made by Charles Gide (Revue Economique Internationale, March, 1910, p. 445):

> Germany increases its population by a million each year, while France increases its capital by two billion francs annually; a million children at 2,000 francs each adds up to two billion.

In this way Gide tries to establish a correlation between the fecundity of French capital and the lack of fecundity in children. He omits the fact that Germany, while increasing her population, was at the same time also accumulating capital at an even faster rate than France.

5. J. Lescure, L'Epargne en France, Paris, 1914, p. 38, passim.

6. R. Pupin, op. cit., p. 127.

7. Pierre Leroy-Beaulieu, Les Impôts et les Revenus (Paris, 1914), pp. 61–62.

national income of the three countries. A. Neymarck in a study on French savings maintained that "It can be proved, in fact, that of these 10 million electors, 9 millions at least have a book at some savings institution, a government 'rente,' a railroad bond, a house, etc."[8] To support his claim that French wealth is widely disseminated, Neymarck indicates that "of 12 million households in France there are 9 million, each possessing its own house."[9] This is a record which cannot be matched by Great Britain or Germany. C. K. Hobson[1] also takes for granted that in France the ownership of property is much more widespread than in the United Kingdom.[2]

Part of this explanation is to be found in the more pronounced agricultural character of France. The typical farmer in France was and is a small landowner, whereas in England and in Germany vast tracts of land were held by large landowners.[3]   In addition, in France the tendency has been towards a decline in the very large ownership of land, as can be seen from Table IX. We note here a definite decline in the number of the very large holdings and an increase in the middle category.

The dispersion of property and income in France appears to

TABLE IX

SIZE OF HOLDINGS IN FRANCE(a)

Size of Holdings	1862	1892	1908
Very small (under 2½ acres)	not reported	2,235,000	2,088,000
Small (2½–25 acres)	2,435,000	2,618,000	2,524,000
Middling (25–100 acres)	636,000	711,000	746,000
Large (100–250 acres)		105,000	118,000
Very large (over 250 acres)	154,000	33,000	29,000

(a) J. H. Clapham, Economic Development of France and Germany (1815-1914), Cambridge, 1921, p. 165.

8. A. Neymarck, National Monetary Commission, Document No. 494 (Washington, 1910), pp. 166 ff.
9. Ibid, p. 173.
1. C. K. Hobson, The Export of Capital (London, 1914), pp. 36-37.
2. This, of course, does not imply that the average income of the Frenchman is higher than that of the German or Englishman. Bresciani-Turoni (Econometrica, 1939, pp. 115-116) observes: "International comparisons also seem to confirm that there is no definite relation between the size of the average income and the degree of inequality of personal incomes. In Italy, both the average income and the inequality of incomes are less than in Great Britain."
3. M. M. Knight and others, Economic History of Europe (Boston, 1928), p. 448, and J. Clapham, op. cit., p. 198.

indicate that the saving process was carried out by the masses of Frenchmen, while in other large Western countries a larger proportion was effected by the more numerous richer classes and by corporations. In our opinion this fact also explains the phenomenon that, in spite of the proportionately far greater number of savers and earners, France did not increase its accumulation of wealth to any greater degree than Germany or England, either in the absolute or proportional sense.

We believe, however, that Lescure's argument emphasizing "a century of continuous accumulation of wealth"[4] merits special attention. This seems to be the only way to account for the phenomenon that on the eve of the World War, France, in spite of lower prevailing wages, and in spite of proportionately smaller annual savings, was estimated to have a per capita wealth position similar to that of Great Britain and superior to that of Germany whose annual savings were considerably above those of France between 1870 and 1914. Josiah Stamp cites the following amounts of national capital per head of population in 1914:[5]

  United Kingdom.................318 pounds
  France........................303 pounds
  Germany.......................244 pounds

Helfferich's estimates[6] of per capita national wealth for the years 1908–1910 are:

  United Kingdom.................5100 to 5800 marks
  France........................5924 marks
  Germany.......................4600 to 4900 marks

Both of these authors made use of various reliable sources in arriving at these amounts, and both agree that the per capita wealth of France before the World War was superior to that of Germany and about on a par with that of the United Kingdom. In view of our detailed discussion on wealth in each of the three countries, the inference is inescapable that France before the 1870's excelled in the per capita accumulation of wealth as compared with England and Germany.[7]

4. J. Lescure, op. cit., p. 40.
5. J. Stamp, Studies in Current Problems in Finance and Government, p. 332.
6. K. Helfferich, op. cit., p. 113.
7. This conclusion is confirmed by R. Pupin's observation that in spite of the very favorable conditions for savings that existed in France during the

The wider dispersion of income in France was one of the factors that led to lower national savings. A relatively more equal distribution of income tends to raise the consumption function of the nation at the expense of savings, since people with a lower average income must spend a larger share of it on current expenditures. On the other hand, the greater inequality of wealth and income in Germany and England was one of the reasons for a higher ratio of savings to national income as compared with France.

Another possible explanation is the greater industrialization of Germany and the United Kingdom. A larger proportion of national wealth was therefore concentrated in business and industrial enterprises in these two countries than in France, with the resulting effect on the flow of the national income. Since industry and commerce were on the increase in Western Europe during that period, larger shares of the profits and consequently of national income were reinvested in the numerous enterprises of that part of the world, especially so in Germany and the United Kingdom. This was in large measure due to the growth of population, which constituted a continually increasing source of demand for capital. As a result more of the national income went into savings ·in Germany and in the United Kingdom.

Keynes[8] ascribes this trend to the spirit of optimism which a growing population engenders in the business community, leading to a rising demand for capital. In his Fiscal Policy and Business Cycles,[9] Hansen describes in detail the effects of a rapidly growing population upon the extensive expansion of investments. He found that the facts of growth and expansion "minimized the risk of new ventures." A German authority on population, August Lösch,[1] while agreeing that an increasing population causes capital to grow proportionately faster, also attributes this tendency to the lesser risk involved in capital expansion which results from population growth, as compared with an expansion "inspired by technical progress or by the gaining of new markets." Whatever the reason, there is agreement among them that the growth of population broadens demand for capital, which acts as an encouragement for increased business and national savings.

years 1885-1910, "we never reached the rate of our grandfathers, who saved 13 per cent of their income" (op. cit., p. 128).

8. J. M. Keynes, Galton Lecture, loc. cit.

9. A. H. Hansen, Fiscal Policy and Business Cycles, 1941, pp. 347-361.

1. August Lösch, this JOURNAL, 1937, p. 658.

A third factor in the explanation is the age distribution. The experience of France during 1870–1911, and that of the other Western countries after the first World War, when they, too, became subject to a declining trend of population, confirm the theory that a stationary population contains a larger percentage of older people. The data in Table X are of interest in this connec-

<div align="center">

TABLE X

Age Distribution of Population
(in percentages)

</div>

	Census	Below 15	15–29	30–59	60 and over
Germany..........	1910	34	26.1	32	7.9
England and Wales..	1911	30.6	26.6	34.8	8
France.............	1911	25.7	23.9	37.8	12.6

tion. They show that France had a substantially higher proportion of people over 60 years of age than either Germany or England. It is also well known that the French traditionally retire from business and occupations at an earlier age than do most other Western peoples. Old and retired persons act almost entirely as dissavers and increase the consumption function of a nation, thus concomitantly decreasing its savings propensity. This is what happened to the French.

On the other side of the ledger, acting as a factor reducing the consumption function of the French, has been the phenomenon of a smaller number of children as compared with Germany and England. Children in modern countries are almost entirely consumers. However, their effect on savings is quite restricted in intensity, since the expenditures per child are smaller than per adult, and since as a rule an only child receives better care. This limitation of the effect of a smaller family on reducing expenditures was recognized by French students of population.[2] On the whole, it is therefore safe to infer that the total effect of the age distribution in France was in the direction of more pressure on expenditure and towards lesser savings.

For the period under study the most accepted classification of productive ages appears to be the years 15 to 59. The French official statistics lay special stress on the ages of men from 15 to 59 as most representative of the productive capacity of the nation.

2. J. Lescure, L'Epargne en France (Paris, 1914), p. 38.

Table XI provides evidence of a consistently higher percentage of such men in stationary France, as compared with England and Germany.

TABLE XI

PERCENTAGE OF MEN OF THE AGES 15–19(a)

	France	Germany	England and Wales
1870.	30.5	23.7(b)	27.1
1880.	29.9	27.6	27.1
1890.	30.5	27.8	27.7
1900.	30.4	28.2	28.9
1910.	30.5	28.7	29.5

(a) Annuaire Statistique de la France, 1932, p. 229.
(b) In 1870 this refers to the age group 20–59.

France appears to have consistently larger numbers of men in these age groups as compared with the other two countries. Table XI shows that in spite of a higher proportion of older ages within a stationary population, its proportion of active ages (15–59) is not lower than that of a growing population. On the eve of the first World War the three countries under discussion had the following percentages of active ages:

$$\text{Germany} \dots\dots\dots\dots\dots\dots\dots\dots\dots 58.1$$
$$\text{England and Wales}\dots\dots\dots\dots\dots\dots 61.4$$
$$\text{France}\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots 61.7$$

All evidence also points in the direction of a proportionately higher available labor force in a stationary population, since in such a state a larger percentage of women is employed in industries, trades and the professions. This fact is to be ascribed primarily to the lesser family responsibilities resulting from a lower birth rate. Table XII shows a large disparity in the proportion of women employed in the three Western countries. This may be explained, in part, by the importance of the textile and clothing industries in France, industries in which, as a rule, women form the bulk of the workers. But this accounts for only part of the increased ratio of employed women. Even in these industries France shows a much larger proportion of women employed. Out of each 100 workers in the textile and clothing industries the following numbers were women:

	France		*Germany*
1896	72.47	1895	53.98
1906	76.67	1907	57.45

TABLE XII

NUMBER OF WOMEN PER 100 OCCUPIED PERSONS(a)

		Professions (Army and Agriculture not Included)	Industry and Commerce Combined
France	1866	36.5	27.2
	1896	36.7	32.5
	1901	37.9	34.0
	1906	38.5	
Germany	1882	29.2	18.4
	1895	29.0	20.2
	1907	27.8	20.9
United Kingdom	1881	35.4	24.4
	1891	35.2	24.3
	1901	32.7	23.0

(a) Bulletin de la statistique générale, Vol. II, p. 181.

It seems plausible to infer that as a result of smaller families or childless marriages, more women drift into remunerative occupations, since their home duties are considerably reduced. In the voluminous study on comparative costs of living in French, German and English towns in the 1900's by the British Board of Trade,[3] it is brought out that wives contributed to the income of French families to a much larger extent than in either Germany or the United Kingdom.

In our opinion, this larger participation of women in gainful employment explains two interesting phenomena: (1) that in spite of higher prevailing wages in Germany and in England,[4] it was revealed by the British Board of Trade Inquiries that 35 per cent of French families earned more than 40 shillings a week, while in England this percentage was 31 per cent, and in Germany only 15 per cent; (2) that in spite of higher wages paid to German workers,

3. Cost of Living in German Towns, Report of an Inquiry by the Board of Trade, London, 1908; Cost of Living in the Principal Towns of the United Kingdom, London, 1908; Cost of Living in French Towns, London, 1909.

4. The British Board of Trade Inquiries, after exhaustive comparative studies of weekly wages in fourteen different occupations, arrived at the following respective ratios: (For 1904-1906) England, 100; Germany, 83; France, 75.

the French per capita income was greater than that of Germany on the eve of the first World War.

Josiah Stamp's estimates of per capita income for 1914 were (in pounds sterling):

United Kingdom...................... 50
France 38
Germany............................. 30

Yet, in spite of a higher average family income, France saved a smaller proportion of her national income than either of the other two countries. This appears to be mainly explained by population factors, some of which were discussed above.

SAVINGS AND FOREIGN INVESTMENTS

France. As mentioned before, France during the decades preceding 1914 differed from Germany and England, not only in the practice of saving, but also profoundly in the scope and type of foreign investments. The stationariness of the French population appears to have led to a lower demand for capital in France, and seems to have been the major single cause for the diversion 'of so large a share of the French capital abroad. This, in spite of the fact that France saved a smaller share of her national income, and that France, during the decades under discussion, had a somewhat lower rate of return from her foreign investments than either England or Germany.[5]

Interesting and instructive information is available concerning French exports of capital. Neymarck[6] gives data on French and foreign securities in France (Table XIII). The important lesson to be learned from these figures is that between 1890 and 1912 the total amount of securities floated in France was 34–41 billions, and that 20–23 billions of this were for investments abroad and for

5. Moulton and McGuire (Germany's Capacity to Pay, New York, 1923, p. 258) quote several authorities on the prevailing returns from foreign investments by different countries during the prewar decades:

G. Paish placed the English returns from foreign investments at 5 per cent.

Riessers estimated the German returns from foreign investments at 6 to 5 per cent.

R. Pupin placed the French returns from foreign investments at 4½ per cent.

6. Neymarck seems to be the undisputed authority in this field in France; all other estimates seen by the author, such as those given by Moulton and Lewis, Lascaux, Schmoller, etc., are borrowed from Neymarck. A. Neymarck. Bulletin de l'Institut International de Statistique, Vol. XX, Part II, p. 1406,

TABLE XIII

AMOUNT OF OUTSTANDING SECURITIES IN FRANCE
(In Billions of Francs)

Year	Amount of Capital in Form of Stocks	Of that Amount, the Foreign Securities (including loans to foreign governments)
End of 1869.......	33	10
End of 1880.......	56	15
End of 1890.......	74	20
End of 1902.......	87–90	25–27
End of 1904.......	90–93	27–30
End of 1906.......	97–100	30–32
End of 1908.......	103–105	32–35
End of 1910.......	106–110	38–40
End of 1912.......	108–115	40–42

foreign loans.[7] Another unique aspect of this situation is that of the 40–42 billion francs invested abroad, 30–33 billions, or three-fourths, consisted of loans granted to foreign governments, while only 10–12 billions went into industrial and commercial foreign stocks and bonds.

Quite a different situation prevailed in Germany and in the United Kingdom, where the bulk of the exported capital was used for the financing of industry, frequently enterprises managed by the emigrants of the two countries. The stationariness of French population, which probably is largely responsible for the absence of French emigration of any importance, may thus seem to have had an effect also upon the type of French investments abroad.

It would, perhaps, be stressing the point too much to say that this was the only reason for the extraordinarily high proportion of French investments in foreign governmental bonds. In her constant fear of Germany, France was led to use her financial resources to acquire allies against her potential enemy. For example, the enormous loans granted to the Russian government before the first World War were primarily of strategic significance. On the other hand, it is fair to assume that the relative absence of French emigrants abroad had a certain influence on the character of French foreign investments.

The total growth of French wealth during 1892–1908 was estimated by Théry at 44 billion, and by Pupin (1893–1911) at 36

7. A. Neymarck, ibid, p. 1410.

billion francs. This means that at least 50 per cent (probably more) of the accumulated French wealth during the two decades preceding the World War went into foreign investments.

In terms of percentage growth, the difference between internal and foreign investments is even more striking. During 1893–1911 total wealth, including foreign investments, grew by 16–17 per cent, while investments abroad grew by 100 per cent. Taking only the growth of wealth within France, we would get an increase of investments in France of only around 9–10 per cent during that period. This is an indication of the tremendous rôle played by foreign investments during the two decades preceding the World War. Between 1878 and 1911, according to Pupin's data, investments within France grew by 28 per cent, while investments abroad for the same period grew by 165 per cent.

Great Britain. Table XIV gives estimates of the capital of Great Britain within and outside the Isles during the years 1875–

TABLE XIV

BRITISH CAPITAL 1875–1909 (a)

Year	Capital in U.K.	Capital Abroad	Total *
1875................	....	....	8,545
1885................	8,735	1,320	10,037
1895................	9,063	1,600	10,663
1905................	11,009	2,025	13,036
1909................	11,654	2,332	13,986

(a) C. K. Hobson, op. cit., p. 207.

1909. According to these figures, a little over one-fourth of the increase in wealth for the years 1885–1909 was invested abroad.[8] These estimates are interesting inasmuch as they show that in Great Britain too foreign investments increased at a greater percentage rate than domestic capital. (For 1885–1909 capital abroad grew by 79 per cent, and capital within the United Kingdom grew by 33 per cent.) It should be kept in mind, however, that Great Britain, with her colonial possessions and her long traditions in the matter of foreign investments, was in a very favorable position with reference to the diversion of capital abroad. Yet it is interesting to note that while the trend towards foreign investments

8. The term "abroad" includes, in this case, the British dominions and colonies.

was high in England, the proportion of her invested capital abroad was far lower than that of stationary France.

Germany. Helfferich, Crammond, and other students of German wealth and income concur in the opinion that in the decades preceding World War I the amounts invested abroad by Germany were substantially less in the later years than those in the earlier years of that period.[9] There are no authoritative estimates of German foreign investments before the 1890's, but from all evidence they appear to have been rather small. Gustav Schmoller places German capital abroad for 1892 at 10 billion marks. Moulton and McGuire quote several other estimates for 1892-1893, all of which are in substantial agreement that the amount of 12 billion marks represented German foreign investments for that year. The same study notes a similar agreement for the year 1914, placing German capital abroad at 25 billion marks.[1] This represents a growth of 13 billion marks during the 21-22 year period.

Our study of the growth of German wealth for the years 1895-1913 reveals an increase of German capital of over 100 billion marks,[2] which leads to the conclusion that not more than 12-13 per cent, and probably closer to 10 per cent, of the net addition to accumulated German wealth was invested abroad during the two decades preceding the World War. These ratios are vastly lower than those for England and France, and serve as eloquent testimony to the rapid tempo of the internal industrial and commercial growth of Germany. Practically all the capital available found profitable outlets within Germany, and a large part of this growth may be explained by the unparalleled increase in her population.

CONCLUSIONS

Our study gives support to the hypothesis advanced by Keynes and Hansen that the outlet for home investment is directly correlated with the population trend, with the result that stationary France invested at home the lowest proportion of her national income. France of 1871-1911, in spite of saving a smaller share of

9. Moulton and McGuire, op. cit., p. 280.
1. In a recent work entitled German Economy, 1870-1940 (New York, 1940), Gustav Stolper estimates the net German foreign investments in 1914 at 25 billion marks.
2. Leon Goldenberg, Income and Savings in France, 1871-1914 (unpublished), p. 139.

her national income than either Germany or Great Britain, was confronted with a relative abundance of capital, more than half of which was forced to seek employment abroad. Furthermore, a considerable part of that half of French savings which went into home investments was absorbed by government issues.[3] The French national debt grew from 1,280,000,000 francs in 1815 to 33,858,000,000 in 1914. A large share of these funds was used for armament expenditures.

This pattern worked quite well before the first World War, and was probably instrumental to an important degree in maintaining the steady upward trend in the per capita national income of France during 1870–1914, in which respect she did not materially differ from Germany and the United Kingdom for the same period. The annual per capita rise in income in all three instances was between 1.1 per cent and 1.4 per cent. There may be some doubt as to whether opportunities for safe foreign investment of excess capital would now be available to all Western countries experiencing stationary or declining trends of population, trends which in recent years have been characteristic of most modern industrial nations.

<div align="right">LEON GOLDENBERG.</div>

WASH NGTON, D. C.
DEPARTMENT OF STATE

3. A. H. Hansen, Full Recovery or Stagnation, 1938. p 324

APPENDIX

TABLE I

PROPORTION OF LIVE BORN CHILDREN PER 100 INHABITANTS[1]

(Territory of 1914)	1871–80	1881–90	1891–1900	1901–10
France.............	2.34	2.39	2.22	2.06
Germany...........	3.91	3.68	3.61	3.30
England and Wales...	3.54	3.25	2.99	2.72

1. Annuaire Statistique de la France, 1932, p. 238.

TABLE II

ANNUAL NUMBER OF DEATHS PER HUNDRED OF POPULATION[1]
(Stillbirths not Included)

(Territory of 1914)	1871–80	1881–90	1891–1900	1901–10
France.............	2.26	2.05	1.92	1.64
Germany...........	2.72	2.51	2.22	1.87
England and Wales...	2.14	1.91	1.82	1.54

1. Annuaire Statistique de la France, 1932, p. 239.

TABLE III

ANNUAL PER CAPITA INCOME IN FRANCE, GERMANY,[1] AND UNITED KINGDOM (1870–1913)

| France | | | Germany | | | United Kingdom | | |
Year	Per Capita Income (in francs)	Authority	Year	Per Capita Income (in marks)	Authority	Year	Per Capita Income (in pounds)	Authority
1870	530	Baxter	1869	372	Mulhall	1867	31.4	Levi
1878	588	Pupin	1870	388	Baxter	1870	28	Baxter
1888	608	DeFoville	1891	498	Reichsamt	1875	36.6	Giffen
1893	627	Moulton-Lewis	1895	480	May	1883	35.8	Giffen
1900–1901	672	Seligman	….	….	Mulhall	1883	35.8	Levi
1900	674	Colson	1896	410	Schmoller	1888	35.3	Mallock
1903	710	Pupin	1896	506	Helfferich	1891	42.3	Bowley
1908	713	Lavergne and Henri	1900	556	Reichsamt	1900–1901	41.7	Seligman
1911	808	Pupin	1901	550	May	1903	41.6	Giffen
1911	909[2]	Pupin	1906	616	Reichsamt	1906	46.9	Crammond
1913	905	Bernonville	1907	658	Reichsamt	1907	44.4	Bowley
1913	907	Colson	1911	698	May	1907	44.9	Mallock
….	….	….	1912	600	Reichsamt	1911	46.1	Stamp
….	….	….	1913	725	Helfferich	1913–14	48.7	Stamp
….	….	….	1913	748	Lascaux	….	….	….
			1913	….	Reichsamt	….	….	….

1. Leon Goldenberg, Income and Savings in France, 1871–1914 (unpublished), p. 108.
2. This estimate includes corporation reserves and cost of food consumed by farmers

TABLE IV

PERCENTAGE GROWTH OF PER CAPITA NATIONAL INCOME[1]

Estimates	France		Germany		United Kingdom	
	Years	Annual Per Capita Percentage Growth	Years	Annual Per Capita Percentage Growth	Years	Annual Per Capita Percentage Growth
I	1870–1911	1.3	1870–1912	1.4	1867–1911	1.1
II	1870–1911	1.7	1870–1911	2.0	1870–1911	1.5
III	1870–1893	0.8	1870–1895	1.0	1867–1891	1.4
IV	1893–1911	2.0	1891–1911	1.9	1888–1911	1.3
V	1900–1911	1.8	1896–1912	2.7	....	....
VI	1903–1911	1.8	1901–1911	2.5	1903–1911	1.2

1. Leon Goldenberg, Income and Savings in France, 1871–1914 (unpublished) p. 108.

ADMISSION POLICIES OF LABOR UNIONS

SUMMARY

The rapid growth of labor unions and the vast extension of collective bargaining in recent years have focused increased attention on the internal affairs of these organizations which wield such great economic power. Numerous charges of dictatorship and racketeering in particular unions have led many to believe that unions are generally undemocratic and irresponsible, but very little attempt has been made to determine the extent to which these practices exist or what is typical of the labor movement as a whole.

One important part of the problem of internal democracy is the admission policies of unions, for the right of a worker to join a union is preliminary to all other parts of the problem. In this particular area there have been frequent charges of racial discrimination, closed memberships, and high initiation fees against individual unions, but no critical study has yet been published to indicate what proportion of unions actually deny admission to workers who desire to join.

The right of a worker to join a union is, in the first instance, governed by the parent or international union, for its constitution and by-laws are binding on all local unions. The international constitution usually limits the right to join by requiring the worker to meet certain qualifications before he is eligible for membership.[1] It may, for example, require him to be employed

1. Nearly all international constitutions prescribe the procedure by which locals shall admit new members. The usual provision is that an eligible worker shall be admitted on a majority vote of the local. The Brotherhood of Locomotive Firemen and Enginemen in its constitution provides that two "blackballs" will exclude, while the Boilermakers provide that three "blackballs" can be overridden only by a two-thirds vote of the local. In a few unions such as the Newspaper Guild, the Typographers union, and the Brick-

within the jurisdiction of the union, prove his skill, not be a member of a rival union, be white and male.[2] The international constitution may also protect the right to join by prohibiting exclusion for certain reasons such as race, creed, nationality, or political beliefs. Within these loose limits, the local union is completely free to accept or reject any worker who seeks to join. Even these loose paper restrictions on the local's power are meaningful only to the extent that the right to join is realized in practice. The international may be unwilling or unable to compel compliance with its constitution and by-laws.

In order to determine the extent to which unions exclude workers from membership, an analysis has been made of the admission provisions in the constitutions of 185 international unions, including practically all internationals of any importance. All available studies relating to admission policies have been consulted in order to get as accurate a picture as possible of the practices of unions apart from their constitutional provisions. In addition, an attempt has been made to state the reasons why unions believe it is necessary to exclude workers within their jurisdiction, and to analyze some of the economic forces which tend to limit or expand the ability of unions to make their exclusionary practices effective.

CONSTITUTIONAL PROVISIONS AND UNION PRACTICES

In the overwhelming majority of cases, unions freely admit all workers within their jurisdiction without any substantial restrictions. It can fairly be stated as the general rule that the right of the worker to join is complete, and that this right can be defined in terms of the exceptions to this general rule. This can be done by determining the extent to which unions exclude persons within their jurisdiction because of (a) race, (b) lack of citizenship, (c) political beliefs, (d) creed, (e) sex, or (f) because the union has closed its membership. These are the principal grounds for exclusion around which any substantial contention has centered.[3]

layers, a rejected candidate may appeal to the National Executive Board, but this privilege is rarely exercised. In practice, the power to admit or exclude lies largely in the membership committee which investigates the applicant, for its recommendations are usually followed without question.

2. See Table I at end of article.

3. There are a few unusual grounds mentioned in various constitutions. The Blacksmiths exclude members of "the state militia, sheriff's office, police force, secret service, or miner's police force." The United Mine Workers

They will be discussed in turn, determining first the extent to which the international constitutions and by-laws exclude otherwise qualified workers, and second the extent to which local unions practice exclusion irrespective of the paper provisions.

Exclusion Because of Race. From its very beginning, the labor movement in this country has wrestled with the problem of admitting Negroes. Many of the early craft unions either excluded them or segregated them in separate locals.[4] Although the American Federation of Labor has repeatedly reaffirmed its desire to include all workers regardless of race or color,[5] it has failed to enforce that policy against a number of its constituent internationals which continue their exclusionary practices.[6] In contrast, the Congress of Industrial Organizations has made non-discrimination one of its basic principles and has countenanced no internationals excluding because of race or color.[7] Because of the autonomous nature of these federations, effective control of admission lies in the international, and it is chiefly their constitutional provisions and established practices which govern.

The constitutions[8] of nine international unions explicitly deny the right to join on racial grounds. Five of these are affiliated with the AF of L[9] and the remaining four are the big independent railroad brotherhoods.[1] Most of these unions provide that a worker to be eligible for membership must be "white," but the Railway Mail Association is more explicit, specifying those of

exclude members of the National Chamber of Commerce, and the Longshoremen (AF of L) exclude anyone "dealing in spirituous liquors." Such provisions are of little practical importance. Many unions exclude foremen or supervisory employees, and many exclude members of rival unions. These provisions are of importance, but they are generally conceded as being entirely proper.

4. Spero and Harris, The Black Worker (1931), Chap. II.

5. Franklin, Negro Labor Unionists in New York (1936), App. II.

6. In the AF of L national convention in 1943, A. Phillip Randolph, president of the Brotherhood of Sleeping Car Porters, vigorously condemned the hypocrisy of the Federation in condoning discrimination practices by many of its internationals. See Report of Proceeding of 63d Annual Convention of AF of L, 1943, p. 421–422.

7. Northrup, Organized Labor and the Negro. (1944), pp. 8 et seq.

8. Statements concerning provisions in union constitutions are summarized in Table I. Membership in the union is indicated by parentheses.

9. Airline Dispatchers, Railroad Telegraphers (30,000), Railway and Steamship Clerks (204,200), Railway Mail Association (21,800), and Switchmen (9,300).

1. Locomotive Engineers (76,000), Locomotive Firemen and Enginemen (119,686), Railroad Trainmen (210,570), and Railway Conductors (36,000).

"Caucasian race or native American Indian." The Firemen's constitution defines "white" as excluding "Mexicans, Indians, or those of Spanish-Mexican extraction." Although the powerful Machinists union has no such provision in its constitution, it obtains the same result by having a provision in its ritual.[2]

These provisions are aimed primarily at Negro workers, for they make up 95 per cent of the non-white laboring force,[3] but it is clear that these provisions by their terms also exclude workers of oriental ancestry, who in some sections of the country constitute a substantial group of workers.

Five large AF of L unions, instead of openly excluding Negroes, provide for admitting them to auxiliary unions.[4] An auxiliary union is formed only where there are enough Negro workers to make a separate local and where a white local already exists. The auxiliary has no voice in national union affairs, is represented by the officers of the white local in all bargaining with the employer, has no share in the grievance procedure, and must depend upon the business agent of the white local for all work assignments.[5] Admission to such an auxiliary must be recognized as another form of exclusion, for the bare "privilege" of paying dues and abiding by union rules can scarcely be said to amount to the right to join.

While a total of fifteen international unions expressly deny the right to join because of race by provisions in their constitutions and by-laws, forty-seven expressly protect the right to join by provisions which prevent exclusion because of race. Twenty-nine CIO, fourteen AF of L, and four independent unions grant

2. In 1890 the AF of L refused to accept the Machinists because of their constitutional provision excluding Negroes. The Machinists shifted it from their constitution to their ritual and were admitted to the AF of L in 1895. Franklin, op. cit., p. 280. The Machinists have a membership of 625,000.
3. Negro Handbook (1944), p. 215.
4. Blacksmiths (10,000), Boilermakers (336,900), Maintenance of Way Employees (116,900), Railway Carmen (95,800), Sheet Metal Workers (25,000). The Railway and Steamship Clerks exclude Negroes by their constitution but regularly admit them to auxiliaries, Northrup, op. cit., p. 3. For a brief history of the Clerks auxiliary unions, see Reynolds and Killingworth, Trade Union Publications (1944), p. 76.
5. For a description of the status of auxiliaries, see Northrup, op. cit., p. 3. It is important to distinguish so far as possible between auxiliaries and segregated locals. The latter have full powers of self-government and are on an equal status with white locals. However, the line between the two is difficult to draw with accuracy, for Negro locals may be subordinate to white locals in varying degrees.

such constitutional protection.[6] Some unions, such as the Wood-workers (CIO), specifically prohibit exclusion by providing in their constitution that "No worker otherwise eligible to member-ship shall be discriminated against or denied membership because of race . . . ," but others, such as the United Mine Workers, simply provide that all workers within the jurisdiction "shall be eligible regardless of color." Although the latter type of provisions might by strict legalistic construction be interpreted as not binding the locals to admit Negroes, they appear to have the practical effect of protecting the Negro's right to join. The Bricklayers (AF of L) protect against exclusion by providing that no person shall be blackballed except for incompetency, and that any discrimination because of race will be punishable by a fine of $100.

These provisions purport to limit the power of the locals to admit or exclude workers, but the local practice may deviate from this norm because the international is either unwilling or unable to compel compliance.[7] Thus the Machinists local at Lockheed-Vega has admitted Negroes in spite of the express prohibition in the international's by-laws,[8] and the Providence local of the Boilermakers has admitted Negroes on a basis of equality, even though the international constitution provides for exclusion by admitting only to auxiliaries.[9] Likewise, the Atlanta local of the UAW (CIO) has refused to admit Negro janitors who were otherwise eligible, in spite of the international's constitutional provision against exclusion and strong policy against discrimination.[1] However, it appears that in the great majority of cases the provisions or policy of the international are voluntarily obeyed by the local unions. In a few cases the international has taken steps to compel recalcitrant locals to comply.[2] Thus the

6. See Table I.

7. Wolfe, Admission to American Trade Unions (1912), p. 126.

8. Northrup, op. cit., p. 207.

9. I Yearbook of American Labor (1945), p. 375. The Boilermakers international has now voted to give auxiliaries more autonomy and repre-sentation in the Metal Trades Council and the National Convention.

1. Northrup, op. cit., p. 191; Weaver, Negro Labor (1946), p. 68.

2 Northrup's analysis of unions in a number of different industries makes fairly clear that the policies within any one international are quite uniform. It seems safe to say that the policy of either denying or protecting the right to join because of race is made on the national level.

National Executive Board of the Bricklayers used a threat of fines to compel the St. Louis local to admit two Negroes.[3]

One hundred twenty-three of the unions surveyed have no provision whatsoever concerning exclusion because of race in their international constitutions. This means that in two-thirds of the unions there is no limitation on the local's power to admit or exclude Negroes as it sees fit. It is essential, therefore, to examine the practices of those unions which have no limiting provisions.[4]

By established custom, nine AF of L[5] and five independent[6] unions refuse to admit Negroes to their locals. In addition, one AF of L[7] and two independent[8] unions apply the more subtle method of relegating Negroes to auxiliaries. Although the individual local unions may be free from constitutional limitations, it is found that they will usually follow the pattern of their parent. There may be some deviation from the pattern within the local unions, such as a few Northern locals of the Plumbers and some West Coast locals of the Electrical Workers admitting Negroes in spite of the strong national policy of exclusion,[9] but such instances are the exception rather than the rule. The exclusionary policies of these seventeen unions are almost as effective and complete in barring Negroes as if there were express provisions in the international constitutions.

On the other hand, many unions have a strong national policy against discrimination because of race. All of the CIO unions have admitted Negroes on an equal basis and have worked hard to break down racial prejudices.[1] Both the United Mine

3. Northrup, op. cit., p. 41.
4. Material on exclusionary practices of unions is taken from Northrup, op. cit., pp. 3–5, and Spero and Harris, op. cit., Chaps. IV and V. Some unions listed by Northrup as excluding Negroes by constitutional provision were found upon investigation to have no such provisions. These unions have been included among those which practice exclusion.
5. Airline Pilots (2,700), Asbestos Workers (4,000), Electrical Workers (312,900), Flint Glass Workers (25,600), Granite Cutters (4,000), Masters, Mates, and Pilots (3,000), Plumbers (130,000), Seafarers (30,000), Wire Weavers (400).
6. Marine Firemen (3,000), Railway Shop Crafts, Railroad Yardmasters of America (3,500), Railroad Yardmasters of North America, Train Dispatchers (3,500).
7. Federation of Rural Letter Carriers (400).
8. Federation of Railroad Workers, Rural Letter Carriers Association (28,066).
9. Northrup, op. cit., p. 211.
1. American Civil Liberties Union, Democracy in Trade Unions (1943), p. 13.

Workers and the Ladies' Garment Workers have repeatedly used pressure to compel locals to admit Negroes.[2] Although the policy of non-discrimination of such unions is violated by some locals, especially in the South, they predominantly follow the pattern of their parent.

In summary, two things should be noted. First, recognition of the right to join rests primarily in the international union. Its constitution and by-laws can bind the local, and its policies apart from the constitutional provisions will almost always be followed by the local. Deviations from the international's norm do exist, but they are the exception. Second, denial of the right to join because of race is not typical of the labor movement, for only a relatively small proportion of unions practice discrimination.[3] Out of 185 international unions, only 32 (20 AF of L and 12 independent) deny the right to join either by their constitutions, their by-laws, or by their established practices. These unions have about 2,500,000 members out of a total of 13,500,000 members in all the unions surveyed.

Although the proportion is small, certainly the extent of discrimination is sufficient to merit serious consideration. Numbers, moreover, do not tell the whole story, for exclusion is not evenly distributed but is concentrated in two main areas. Negroes are barred from almost all of the unions in the railroads with the exception of the Redcaps and the Sleeping Car Porters. They are also barred from a large portion of the skilled craft unions, such as the Boilermakers, Machinists, Electrical Workers, and Plumbers. It is to be noted that these are some of the most powerful unions, and that they control many of the higher paying and more desirable lines of endeavor.[4] In view of this, the problem becomes one of pressing importance.

Exclusion Because of Citizenship. The vast influx of immigrants, in some years going well over the million mark and totaling

2. Northrup, op. cit., p. 161.

3. Apparently the extent of union discrimination because of race is exceedingly small compared with the extent of employer discrimination. In the last six months of 1943, the FEPC had only 82 complaints against labor unions as compared with 1848 complaints against employers. Labor Fact Book No. 7 (1945), p. 171. In its first year of operation, the New York Anti-Discrimination Committee had only 15 complaints against unions as compared with 205 complaints against employers. Annual Report (1946), p. 19.

4. Northrup, op. cit., pp. 6–34, 49–100, 211–217, and 233, 234.

38 million in a period of 100 years,[5] has always been viewed with apprehension by the American worker. Employers sought these newcomers as a source of cheap labor, used them as strikebreakers, and hoped that the Old World antagonisms would keep the workers divided. In spite of the fact that unions were built predominantly from immigrant workers,[6] the American Federation of Labor as early as 1897 urged limitation of immigration. Its continuous efforts[7] culminated in the National Origins Act of 1924, which limited the yearly quota to 150,000.

It was inevitable that this deep antagonism to immigrant labor should be reflected in the admission policies of national unions; and although the then compelling reasons are now gone,[8] the provisions in the constitutions and the union practices still persist.

The constitutional provisions vary widely in their discrimination against aliens. Seven small unions require full citizenship as a qualification of membership, thus completely barring all aliens.[9] Twenty-two unions, including such large ones as the Carpenters, Hodcarriers, and the Teamsters, require that the worker must have filed a declaration of intent to become a citizen.[1] Four of these, the Brewery Workers, Building Service Employees, Hotel and Restaurant Employees, and the Musicians have the additional requirement that the worker become a citizen as soon as he is eligible. The Marine Cooks require only that the worker be "eligible for citizenship."[2]

Some unions, while not explicitly limiting membership to citizens, achieve the same effect by charging aliens special initi-

5. Daugherty, Labor Problems in American Industry (1941), p. 243.

6. Hourwich, Immigration and Labor (1912). In Chap. 15 he presents evidence to show that the labor movement was essentially of foreign origin and that immigrants joined more readily and were more loyal than native Americans.

7. Lorwin, American Federation of Labor (1933), pp. 53, 272, 402.

8. Immigration since 1930 has been less than 50,000 a year. Field, The Refugee in the United States (1936), p. 20. This excludes the immigration of Mexicans and Canadians, who do not come under the quota system.

9. AF of L: Bookbinders (28,900), Broom and Whisk Makers (300) Glass Workers (25,600), and Glove Workers (3,100). CIO: Marine Engineer, (11,000). Independent: United Licensed Officers (2,000), and Welders (3,000).

1. See Table I.

2. Affiliated with the CIO, membership 6,000. This is in effect an exclusion of immigrants from the Orient, and therefore is an exclusion based partially on race.

ation fees. Although this practice is not as widespread as formerly,[3] five small unions still have such discriminatory provisions.[4] The extreme case is the Wire Weavers, who charge foreign workers a special initiation fee of $1,000.

It is almost impossible to determine the extent to which the practices of local unions conform to the national policy. A survey in New York City found that the Musicians strictly enforced the requirement of becoming a citizen as soon as possible by stamping the word "alien" on the membership card of noncitizens. The Bricklayers go beyond the constitutional requirements and refuse to admit any alien unless he is a skilled worker. In contrast, the Bakery Workers admit aliens without any restriction, even though the constitution requires a declaration of intent; and the Long-shoremen give the business agent power to admit or reject aliens as he sees fit.[5] In spite of these variations, it is believed that the locals usually follow the constitutions fairly closely.

In measuring the denial of the right to join because of citizenship, some summaries may be helpful. Thirty-one unions (27 AF of L, one CIO, and three independents) with a total membership of about 2,500,000 have some form of citizenship bar, but only seven of these with a total membership of about 74,000 bar all aliens. As to the size of the group excluded, there were 2,322,000 aliens gainfully employed in 1940,[6] and two-thirds of all aliens in the country did not have their first papers.[7] Although some of them may have only themselves to blame for not being able to qualify for membership in a union, it is estimated that 90 per cent

3. Wolfe, Admission to American Trade Unions (1912), pp. 102–105.
4. The Glass Cutters (1,600), and the Photo Engravers (10,900), charge a special initiation fee of $200, if the trade is learned in a foreign country, and the Wall Paper Craftsmen (2,800) charge $300. The Stone Cutters (1,900) require that an applicant from a foreign country must pay $100. Wire Weavers (400). All of these are affiliated with the AF of L.
5. From a report by Milton Derber to the American Civil Liberties Union (1943).
6. Handbook of Labor Statistics (1941), p. 298.
7.

Status	All Aliens	All Aliens over 21	Male Aliens over 21
Filed Declaration................	924,524	910,416	574,296
Filed no Declaration.............	2,555,128	2,424,976	942,855
Total			
	3,479,652	3,335,392	1,517,151

From The 16th Census of the United States (1940) Population, Vol. II, p. 10.

are unable for the time to become citizens,[8] and under the Oriental Exclusion Act many of them are not even eligible for citizenship.

Thirty-four unions provide in their constitutions that there shall be no exclusion because of "nationality."[9] Such a provision probably does not protect aliens from exclusion, for the Broom and Whisk Makers have such a provision and at the same time require citizenship as a qualification for membership.

Exclusion Because of Political Affiliations or Beliefs. To those who have been saturated with the polemics against labor unions because of their alleged communistic attitudes it may come as quite a surprise to find that at least 30 international unions, with a total of nearly 4,000,000 members, have constitutional provisions denying the right to join because of political affiliations or beliefs.[1] All of these provisions, with one exception,[2] contain clauses which either expressly or impliedly exclude communists.

Most of these provisions make membership in a forbidden organization the test. Thus the Woodworkers exclude "members of the Communist, Nazi, and Fascist parties." The United Mine Workers add to this list members of the IWW and the Ku Klux Klan, and the Painters add members of the German-American Bund. Some unions, such as the Rubber Workers, include a general clause excluding members of "any other organization which has for its purpose the overthrow of our democratic form of government".[3]

A few of these provisions make personal political beliefs the test. Thus the Bill Posters exclude "anyone advocating overthrow of the government by force"; and the Teamsters bar any member of the Communist Party or anyone who subscribes to its doctrines.[4]

In local union practice, these paper provisions are very rarely

8. Field, "Aliens and Unemployment," this JOURNAL, Vol. 49, p. 533.

9. See Table I.

1. Twenty-three are affiliated with the AF of L, four with the CIO, and three are independent. See Table I.

2. The Blacksmiths exclude only members of the IWW.

3. The Bridge and Structural Iron Workers exclude members of "any organization of Communists," the National Farm Labor Union excludes "Communists, Fascists, Ku Klux Klansmen, and other subversive elements," and the Aluminum Workers (CIO) exclude "anyone affiliated with any organization whose principles and philosophy are contrary to the constitution of the international."

4. The American Editorial Association excludes anyone "espousing Communism, Fascism, Nazism, or any other alienism or advocating the violent overthrow of our national institutions."

put into effect in denying admission to workers who seek to join.[5] Since membership in such organizations is usually not openly avowed, it is seldom known and may be extremely difficult to discover. Personal beliefs may be even more difficult to prove. However, it must be recognized that these clauses, because of their generality, place an almost unlimited power in the hands of the membership committee to exclude anyone of a different political opinion.[6] This potentiality is made clear by the frequent use of these clauses to expel members of rival factions when there is a struggle for power within the union.[7]

Although denial of the right to join on political grounds appears on its face to be quantitatively greater than denial of the right to join because of race or citizenship, it is not in practice as serious a problem. It is of concern primarily because the large number of discriminating clauses offer a considerable potential danger.

In contrast to those unions which deny the right to join because of particular political affiliations or beliefs, twenty-nine unions protect the right to join by constitutional provisions which prohibit local unions from discriminating on political grounds.[8]

Exclusion Because of Creed. Although a considerable number of unions exclude because of race, citizenship, and political beliefs, only two international unions have been found which have any constitutional restrictions on admission because of creed. The Master, Mates, and Pilots require a worker to be a "firm believer in God, the Creator of the Universe," and the Railway Carmen exclude a worker unless he "believes in the existence of a Supreme Being."[9] There is no evidence that the unions make any attempt to enforce these vague provisions.

There is some reason to believe that isolated locals discrimi-

5. American Civil Liberties Union, op. cit., p. 16.
6. "Communists and Unions," Nation, Vol. 152, p. 228.
7. Schneider, The Workers (Communist) Party and American Trade Unions (1927) is a history of the struggle within unions and shows the free use of the expulsion power to eliminate opposition. In some cases the opponents were expelled on political grounds, even though there was no such provision in the constitution. See also Millis and Montgomery, Organized Labor (1945), p. 179; Perlman and Taft, History of Labor (1933), ch. 40; Lorwin, American Federation of Labor (1933), p. 259.
8. See Table I.
9. Until recently the Wire Weavers admitted only "white Christians." Northrup, op. cit., p. 261, n. 4. This provision no longer appears in the constitution.

nate to some degree against Catholics or Jews, but it is difficult to find any evidence of specific instances of such discrimination. It is well known that some locals in the needle trades in New York are exclusively Jewish, but this is to expedite union business by segregation in order that meetings may be conducted in Yiddish, since many of the workers are unable to readily understand English. Just as many unions have constitutional provisions protecting workers against exclusion because of race, nationality, or political beliefs,[1] forty-six unions protect against exclusion because of creed.[2]

Exclusion Because of Sex. Labor unions in their early day were looked upon, in part, as social clubs for the working man, and in such an organization women had no place.[3] As unions became increasingly interested in the economic welfare of their members, the desire to exclude women continued, because of fear of competition for jobs.[4] The Knights of Labor attempted to organize women on an equal basis, but failed because of the continued opposition of the men.[5] The AF of L also advocated the admission of women and expended great amounts of funds and energy to organize them;[6] but as in the case of Negroes, it was never able to enforce its policy against the international unions. Only gradually have the internationals removed their constitutional bars against women,[7] the last being the United Mine Workers in 1942[8] and the Boilermakers in 1944.[9] It was estimated that about 3,500,000 women belonged to unions in 1944, or about 20 per cent of all women employed in non-agricultural work.[1]

Only eight unions have been found which now have constitutional provisions excluding women, and all but three of these

1. The same clause usually gives protection on a number of different grounds. Thus the National Maritime Union constitution provides that "No person shall be excluded by reason of race, color, religious beliefs, sex, or political affiliation," and the United Electrical Workers (CIO) constitution provides that a worker shall be eligible "regardless of skill, age, sex, nationality, color, religious or political beliefs or affiliations."
2. See Table I.
3. Wolfson, The Woman and the Trade Union (1925), p. 55.
4. Wolfe, Admission to American Trade Unions (1912), Chap. IV.
5. Wolfson, op. cit., pp. 62 et seq.
6. Lorwin, op. cit., p. 108.
7. Wolfson, op. cit., pp. 74 et seq.
8. American Civil Liberties Union, op. cit., p. 16.
9. Labor Fact Book No. 7 (1945), p. 70.
1. Idem.

are railroad unions.[2] The customary provision simply requires a worker to be "male" to be eligible for membership. These unions cover a relatively small number of workers and occupations, but it is believed that the exclusion of women is much more widespread than the constitutional provisions would indicate, especially in the skilled trades.[3]

Here again, many unions take a contrasting position by giving constitutional protection to women in their right to join. The constitutions of 19 unions provide that no worker shall be excluded because of sex, and 19 others provide that "male and female" workers shall be eligible.[4]

It is generally thought that women are excluded only from unions covering those trades in which they are rarely employed. War experience, however, has shown that there are relatively few jobs which women are unable to perform, and the fact that they are not employed in a trade may be the result of an exclusionary policy of the unions.[5] In 1944 there were an estimated 19,000,000 women employed, 85 per cent of whom planned to continue working after the war.[6] For a group of unions, even though the number is relatively small, to exclude 30 per cent of the working population from membership must be considered a substantial restriction on the right to join.

Closed Unions. Workers may be denied the right to join, not because they are members of some particular group which is discriminated against, but because the union has closed its membership books and refuses to accept any new members, or accepts only a few favored ones such as relatives of members.[7]

2. AF of L: Airline Dispatchers, Operating Engineers (100,000), Railway Mail Association (21,800), Switchmen (9,300), Wire Weavers (400). Independent: Railroad Trainmen (210,570), Railroad Yardmasters (3,500), Railway Conductors (36,000).

3. "But more than 25 national unions still exclude women. About half of these are in the railroad industry, and the rest are chiefly in the skilled crafts where women are rarely employed." American Civil Liberties Union, op. cit., p. 16.

4. See Table I.

5. Among the industries which women entered during the war were blast furnaces, rolling mills, machine tools manufacture, and shipyards. Yearbook of American Labor (1945), p. 416.

6. Labor Fact Book No. 7 (1945), pp. 163–164.

7. It is said of the building trades in St. Louis, "A boy has as good a chance to get into West Point as into the building trades unless his father or his uncle is a building craftsman." Apprenticeship in Building Construction, Bureau of Labor Statistics Bull. No. 459 (1928), p. 9.

Only six internationals have provisions in their constitutions relating to this practice. Both the Hosiery Workers and the Horseshoers provide that no new members shall be accepted by a local when any of the local's members are unemployed, and the Brewery Workers provide that no new members shall be admitted unless jobs are available. On the other hand, the Hatters and the Wall Paper Craftsmen provide in their constitutions that the membership books shall not be closed without the consent of the Executive Board of the international, and the Musicians provide that a local must accept all competent musicians. In all other unions the control of the practice lies exclusively in the local union.

It is impossible to determine precisely to what extent the various 75,000 local unions close their membership books, for no systematic study has yet been made. A few horrible examples, such as Local 110 of the Motion Picture Operators in Chicago refusing to accept any new members for 15 years, have been widely publicized,[8] but it is generally agreed that there are relatively few unions which engage in this practice.[9] In periods of normal business activity it is confined mainly to the building trades, motion picture and theatrical fields, building service, food and newspaper delivery, photo-engraving, lithography, bill posting and diamond cutting. In periods of depression and mass unemployment it may appear in many other industries, such as the garment trades, mining, printing, and longshore work.[1] Closed unions are generally more prevalent in the skilled than the unskilled trades, and are limited mainly to the highly unionized trades.

Limitation of membership is often forbidden by the national unions, which naturally prefer an expanded membership which brings in more dues and more power to control wage standards.[2] The Operating Engineers have taken over local unions that refused to accept new members,[3] but more often the parent is either unable or unwilling to enforce its policy against a stubborn local. The closed union may also be forbidden by the collective bargaining

8. Seidman, Labor Czars (1938), p. 174; Toner, Closed Shop (1942), p. 164.
9. Millis and Montgomery, op. cit., p. 260.
1. American Civil Liberties Union, op. cit. (1943), p. 19; Loft, The Printing Trades (1944), p. 213; Lorwin, op. cit., p. 511; Slichter, Union Policies and Industrial Management, pp. 64–67.
2. American Civil Liberties Union, op. cit., p. 19.
3. Slichter, op. cit., p. 66.

contract with the employers. The contracts of the Amalgamated
Clothing Workers in Chicago and Milwaukee have provided for
open unions, and Actor's Equity made an agreement in 1924 with
the Managers Protective Association for a 25-year period which
provided that the union would be kept open and the initiation
fees would not be raised without consent of the Association.[4] This
type of provision is rare, for the employer is either not concerned
about the worker's right to join or is unable to extract this con-
cession from the union.

Although closed unions are the exception and not the rule, the
example of Local No. 3 of the Electrical Workers in New York has
shown that in some trades and some areas they are able to obtain
complete control of the labor market.[5] They have a natural tend-
ency to increase during depression years when jobs are scarce, and
there is some reason to fear that, as unions become more powerful,
denying the right to join by closing the membership books may
become more prevalent.[6]

EXCLUSIONARY DEVICES

Denial of the right to join does not always reveal itself in
express form, but is often achieved by more subtle means. In the
foregoing sections the extent of denial has been surveyed regardless
of how achieved. This must now be supplemented by considering
the various devices which are used within the local unions either to
close the membership or to discriminate against unwanted groups.

Since most unions admit only upon a majority vote of the
membership, or provide for blackballing by a small number,
denial of the right to join may be effected simply by a negative
vote. More specifically designed and generally applicable devices
are high initiation fees, apprenticeship rules, skill tests, licensing
laws, and permit cards.

Initiation Fees. Where locals are given freedom to set initiation
fees, they may set them so high as to effect complete exclusion.
The Motion Picture Operators of Cleveland charge $1,000, the
Seattle Truckers $500, and the Chicago Glaziers $1,500.[7] It is

4. Idem, p. 73.
5. See Allen Bradley v. Local No. 3, 325 U. S. 797, 65 S. Ct. 1533 (1945).
Sullivan, The Labor Union Racket, pp. 21–28.
6. American Civil Liberties Union, op. cit., p. 20; Tipton, "Closed Unions
in Modern America" (1942), Social Science, Vol. 17, p. 81.
7. Arnold, "Labor's Hidden Hold-up Men," Reader's Digest, Vol. 38,
p. 136. The New York Checker's local of the Longshoremen's Union charges

admitted that no one ever pays these fees — they are intended solely to exclude.[8] The fee may be set so high as to discourage many from joining, yet not so high as to exclude everyone. The Perth-Amboy Electricians in 1926 charged $200, and the New York Lathers in 1930 charged $280; but exclusion was not intended to be complete, for there were provisions for paying the fee by installments.[9] This kind of fee serves the dual purpose of partly limiting membership and also enriching the union treasury. In these unions there is a right to join — for a price.

In order to prevent locals from using high initiation fees as a device for closing the membership or substantially discouraging workers from joining, many international constitutions either fix the initiation fee or provide for a maximum which the local can charge. Ninety-eight unions have such provisions in their constitutions. In 88 the maximum is set at $100 or less, and in 71 it is $25 or less.[1] In the other 87 internationals, the locals are free to set the initiation fees as high as they choose. However, a

a $500 fee in advance; the Motion Picture Machine Operators locals in New York and Chicago have levied fees ranging from $300 to $1,000. Local 644 Photographers of the Motion Picture Industry affiliated with the Stage Employees has required $500 on application and $500 on admission. Other locals with initiation fees from $200 to $500 have been the Cement Masons, Motion Picture Studio Mechanics, the Bill Posters, and the Carpet and Linoleum Layers — all of New York City; the Glaziers in Cincinnati, the Electrical Workers in Cleveland, and the Chicago Flat Janitors. American Civil Liberties Union, op. cit., p. 23.

8. William Green, in answer to charges, denied that anyone ever paid the Glaziers $1,500. He said it was purely an exclusionary device. New York Times, May 13, 1941, p. 17.

9. Slichter, op. cit., p. 10.

1.

Affiliation	$25 or less	$100 or less	Over $100	Miscellaneous (b)
AF of L............	31	13	1 (a)	5
CIO..............	24	1	..	2
Independent.......	16	3	..	2
	—	—	—	—
	71	17	1	9

(a) The Elevator Constructors set a fee of $300.

(b) Five of these provide that the fee or the maximum shall be set by the parent. The Hatters provide that the maximum shall be the prevailing weekly wage scale, the Plasterers set the maximum at 10 times the daily wage, the Sheet Metal Workers set it at 100 hours of labor at the minimum wage, and the Stage Employees set it at four times the highest regular weekly wage scale in the union. The last three set a maximum which is probably in excess of $100.

The Operating Engineers have enforced their constitutional limitation by taking over locals which insist on excluding by charging high initiation fees. Slichter, op. cit., p. 66.

study of 300 locals revealed that the high initiation fees occurred primarily in the skilled crafts, but that relatively few locals charge exorbitant fees and that not many workers were affected. In general, the fees in the CIO are considerably lower than those in the AF of L.[2]

It has already been pointed out that initiation fees are used by five unions to discriminate against aliens by charging them a higher initiation fee. A more subtle but more widespread discrimination against Negroes may also be effected; for where the fee is as much as $100, it will exclude many Negroes because they too often lack the financial resources to pay that amount.[3]

Apprenticeship Requirements. Although 51 unions have constitutional provisions regulating apprenticeship, only 21 of these purport to require apprenticeship or working experience for admission to membership,[4] and it is doubtful if these requirements are very rigidly enforced.[5]

These regulations have purposes other than restricting membership. They are designed to guard against employers who would try to undercut union standards by employing large numbers of apprentices at low wages for long terms. They may also be used to limit the number who are trained in a trade in order to prevent an oversupply of workers. At one time the Printers so limited the number of apprentices that fewer were indentured than withdrew from the trade.[6] However, union regulations are seldom the cause of a shortage of skilled workers, for investigations have shown that employers have almost uniformly indentured less apprentices than the regulations would permit.[7]

Requiring apprenticeship or working experience for admission to the union is effective only where the union has control of the opportunities for learning the trade. Technological changes and the increased opportunity to learn trades in schools and by other means have reduced the effectiveness of this device in excluding workers from the union.[8] However, these regulations do offer an

2. Taft, Dues and Initiation Fees in Labor Unions, this JOURNAL, February, 1946, p. 219. (Out of the 300 locals, only 27 charged more than $100 initiation fee, and 205 charged $50 or less.)
3. Franklin, op. cit., p. 274.
4. See Table I.
5. Slichter, op. cit., p. 10; Millis and Montgomery, op. cit., p. 447.
6. Loft, op. cit., p. 214.
7. Slichter, op. cit., pp. 28, 33.
8. Motley, Apprenticeship in American Trade Unions, p. 58; and Wolfe,

indirect but real handicap to the Negro. In the South he rarely has a chance for training in a trade through the schools, and very few white craftsmen will take him as an apprentice.[9]

Competency Tests. Twenty international unions have constitutional provisions requiring competency in the trade as a condition of membership.[1] The Bricklayers require an applicant to pass a skill test; the Carpenters provide that the applicant must be able to command the standard union scale; and the Printing Pressmen require that he be a "qualified worker." These requirements may be perfectly proper to enable a union to guarantee an employer satisfactory workmanship. However, they may be used as a device to exclude any or all applicants as the officers decide,[2] and they have frequently been used as a subterfuge for denying Negroes admission.[3]

License Laws. Many unions in the skilled trades have used political pressure to obtain passage of state laws or city ordinances providing for licensing of workers.[4] Licenses may be required for plumbers, electricians, stationary engineers, motion picture operators, barbers, and many others. The majority of the licensing committee are almost always union members, who are then in a position to exclude unwanted persons from the trade. The Plumbers and Electricians (AF of L) have openly advocated these laws as devices for restricting the number of available workers and especially for excluding Negroes.[5] Inspection laws may be used in the same fashion; for when inspection is controlled by the union, refusal to approve work done can quickly drive non-union workers

op. cit., Chap. II show the decline of apprenticeship. In the last ten years the Federal Government has attempted to revive and stimulate the apprenticeship system. See Federal Committee on Apprentice Training, What the Federal-State Apprentice Training Means to Employers (1935); same, The Apprentice and the School (1939); same, Out of Crisis, Opportunity! (1940).

9. Apprenticeship requirements have tended to eliminate Negroes from the bricklaying trade among others. Northrup, op. cit., p. 38.

1. See Table I. Two unions, the Cannery Workers and the Packing House Workers, provide in their constitutions that a worker shall be eligible regardless of skill.

2. Wolfe, op. cit., p. 70.

3. Franklin, op. cit., p. 274.

4. Slichter, op. cit., p. 47.

5. Negro plumbers in Baltimore were failed because of a deficiency in calculus. Northrup, op. cit., pp. 23, 25. The Locomotive Firemen and Enginemen advocated a federal licensing law as a device for excluding Negroes. Spero and Harris, op. cit., p. 477.

from the trade.[6] Although these laws do not directly exclude workers from membership in the union, they are used to achieve the same end as other exclusionary devices.

Permit Cards. If a union restricts its membership below that which is required to meet the peak demands for labor, it must make some provision for permitting non-union workers to work with union members during the rush period.[7] This is usually done by issuing permit cards for a fee of $1 or $2 a week, which entitles them to work temporarily but gives them no rights as union members. This device has been used by the Carpenters, Electricians, Longshoremen, and Brewers unions to handle peak work loads.[8] It has recently been adopted by the Machinists to allow Negroes to work where the union has a closed shop contract without admitting them to membership.[9]

The permit system may be abused by a union which attempts to capitalize on its monopoly position by charging high fees for the cards.[1] In a few cases it is used as a source of graft by corrupt union leaders, who sell permit cards to outsiders even when regular members are without work.[2] Just as national unions have opposed closed unions, they have also opposed the permit system. A few, such as the Operating Engineers, prohibit locals from issuing permit cards.[3]

WHY DO UNIONS DENY THE RIGHT TO JOIN?

Unions which deny the right to join do not act out of sheer malice or arbitrariness, but upon what seem to them good and sufficient reasons. For the most part these reasons grow out of conditions in our social and economic systems which are beyond the power of the union to control.

First and foremost of the reasons for excluding other workers

6. Northrup, op. cit., p. 23.
7. Toner, op. cit., p. 165.
8. American Civil Liberties Union, op. cit., p. 23.
9. Weaver, op. cit., p. 226.
1. Locals of the Electricians (AF of L) sold permits for $1 or $2 a day during the construction of some of the army camps. Arnold, "Labor's Hidden Hold-up Men," Reader's Digest, Vol. 38, p. 136.
2. Malloy of Local 110 of the Motion Picture Machine Operators in Chicago sold permits for $450 to $1,000 while members of his own union were idle. Seidman, op. cit., p. 174. Kaplan of Local 306 in New York used the same technique to exploit workers. Sullivan, This Labor Union Racket (1936), p. 56. See also Millis and Montgomery, op. cit., p. 260.
3. Slichter, op. cit., p. 70.

is the desire of union members to obtain for themselves a priority on available jobs. With the exception of the war periods, our economy has at no time in the last forty years supplied enough jobs for all who wanted to work;[4] and many of the jobs which it did supply paid extremely low wages.[5] There is, therefore, a constant competition among workers for the available jobs, and there is even keener competition for jobs in those industries where wages are comparatively high. Threatened by these sources of job competition, those who are already members of a union may feel compelled to combine exclusionary practices with the closed shop in order to preserve for themselves the limited number of higher paying jobs which are available.[6]

When technological changes have reduced the number of workers needed, such as happened in the glass industry, the unions have often attempted to counterbalance the decreased demand by closing their membership.[7] During depression years when jobs are scarce, many unions, such as those in the garment trades and mining, close their membership in order to give their older members priority in any work which may be available.[8] In industries where the number of workers required fluctuates, such as in longshore work and the building trades, unions often limit membership to the number who can be regularly employed, thus insuring their members nearly full-time employment.[9] Sufficient temporary permit cards can be issued to supply the needs of the rush periods.[1]

4. Between 1890 and 1940 there have been between 1,000,000 and 15,500,000 unemployed. This number varies between five and 31 per cent of the working force, with an average of about 12 per cent. Daugherty, op. cit., pp. 65-66; Douglas, The Problem of Unemployment (1931), p. 26. In October, 1941, there were still 3,900,000 unemployed. Handbook of Labor Statistics (1941), p. 183.

5. In 1939 over 14 million workers earned less than $600 a year and 22 million earned less than $1,000 a year. Statistical Abstract (1945), p. 409. Farm labor averaged from $25 a month in 1933 to $43 a month in 1941. Idem, p. 153. In 1944 there were four million industrial workers earning less that 40 cents an hour. Labor Fact Book No. 7 (1945), p. 133.

6. Tipton, op. cit., p. 81.

7. Slichter, op. cit., p. 67. The Lacemakers restricted membership between 1907 and 1914 upon the increased use of machinery. Ibid.

8. American Civil Liberties Union, op. cit., p. 21.

9. American Civil Liberties Union, op. cit., p. 21.

1. "If a portion of these permit card men were admitted to regular membership it would create an unwholesome layoff during the dull season under our rotation system, and it is natural that the regular members oppose admitting permit men to regular membership in order to protect themselves against unwholesome layoffs." President of Brewers, Proceedings of 28th

In some cases the use of the closed union has been abused by limiting membership to less than the normal needs of the industry. This practice, which greatly increases the members' earnings through extra pay for overtime work, has been used frequently in the building trades, especially during the construction of army camps at the beginning of the war.[2] It has occurred also in the brewing industry, printing, photo-engraving, theatrical crafts, and among the glaziers.[3]

The closed union is used by workers who seek to eliminate all competition for their jobs. However, they may seek only to reduce competition by excluding particular groups which they believe constitute a special threat to their job security. Thus exclusion of Negroes,[4] aliens,[5] and women[6] has been motivated primarily by a desire to eliminate these particular sources of notoriously cheap labor as potential job takers.

The second reason for denying workers the right to join is the desire to protect the union from internal disruption. The union can be effective in bargaining with the employer and enforcing its demands only to the extent that it can maintain its strength and take unified action. Many unions have refused to admit Negroes because of the fear that it would either drive white workers into unions which excluded Negroes or create such dissension among the members that effective and cohesive action would be impossible.[7] This fear is not totally unfounded, for in some instances

Convention (1942), pp. 132–134, quoted in Shister, Trade Union Government — A Formal Analysis, this JOURNAL, November, 1945, pp. 78, 94.

2. The Steamfitters at Fort Snead admitted only six men so they could obtain large amounts of overtime pay. Arnold, "Labor's Hidden Hold-up Men," Reader's Digest, Vol. 38, p. 136.

3. American Civil Liberties Union, op. cit., p. 21.

4. Northrup, op. cit., p. 5 (most exclusions occur in skilled craft unions which are work-scarcity conscious); Spero and Harris, op. cit., p. 56 ("It is the desire to restrict competition so as to safeguard job monopoly"); Wolfe, op. cit., p. 133 ("The economic is the strongest single factor").

5. Wolfe, op. cit., Chap. VI. Much of the discriminatory policy was aimed not only at excluding aliens who had already arrived but to discourage others from coming. See also Hourwich, op. cit., p. 45; Report of the Immigration Committee, 61 Cong. 3d Sess. Sen. Doc. No. 764 (1911), pp. 371 et seq.

6. Wolfe, op. cit , Chap. IV; Lorwin, op. cit., p. 220.

7. In many instances employers have played upon the race issue in order to create dissension among the strikers. In the coal mines, employers attempted to keep a "judicious mixture" of Negroes and different nationalities in order to prevent unionization. Cayton and Mitchell, Black Workers and the New Unions (1939), pp. 320 et seq.; Northrup, op. cit., p. 161.

admission of Negroes has resulted in weakening the union.[8] However, this danger can be overcome, and it has been overcome by many unions in both the North and the South.[9]

Similar arguments were at one time presented to justify the exclusion of women. It was said that they would destroy the fellowship and spirit of the union meetings, and that they would even cause dissension among the men who would vie for their favor.[1] Needless to say, those supposed fears have proved unfounded.

Exclusion of communists is based almost entirely on the desire to prevent destructive dissension in the ranks. In many cases where communists have obtained a foothold, the union has been torn by struggles between rival factions.[2] This disruption has often resulted in wildcat strikes,[3] loss of bargaining power,[4] interference with the conduct of authorized strikes,[5] and finally a falling off of union membership.[6]

The third, and perhaps the least important, of the reasons for exclusionary policies is that those policies are simply a reflection of the social attitudes of the union members. There is no doubt

8. Spero and Harris, op. cit., pp. 68 et seq.

9. Franklin, op. cit., Chap. VI; Cayton and Mitchell, op. cit., Chap. XVIII.

1. Wolfson, The Woman Worker and the Trade Union (1926).

2. Millis and Montgomery, op. cit., pp. 241–242; Taft, "New Unionism in the United States," American Economic Revue, Vol. 29, pp. 313, 320 (description of the turmoil in the UAW (CIO) and the National Maritime Union because of the struggle for power between the communists and the non-communists).

3. Schneider, The Workers (Communist) Party and American Trade Unions, p. 46, reports that communist-dominated locals of the United Mine Workers in Kansas and Illinois called repeated strikes in order to embarrass Lewis and seize control.

4. Schneider, op. cit., pp. 75, 93, states that the communist faction of the ILGWU during negotiations caused numerous disturbances, riots, and wild rumors. They even warned employers not to make a contract with the authorized representatives of the union.

5. Schneider, op. cit., pp. 100–104. The communist faction led the industry-wide strike of the ILGWU in 1926, and though they had the full support of the union, they failed to win the strike because they attempted to use the strike for political ends.

6. Among the many unions which at one time or another have been weakened by internal struggle with the communists are the Machinists, United Mine Workers, Amalgamated Clothing Workers, Furriers Union, and ILGWU (Schneider, op. cit., pp. 23, 50–59, 67, 74, 104) and the American Federation of Teachers (Nation, Vol. 152, p. 307.) See also Lorwin, op. cit., p. 259, and Perlman and Taft, op. cit., Chap. 40.

that exclusion of Negroes is based partly on social prejudices.[7] This is especially true in the South, where admitting them may seem to violate many of the deep-rooted social practices. However, the appeal to prejudice is probably mainly the emotionalizing of an underlying desire to protect an economic advantage, with the Negro the most likely victim. In many instances aggressive unions have been able to break down some of the prejudices of their members,[8] even to the extent of calling strikes to prevent discrimnation by the employer.[9]

Exclusion of aliens was in part based on a dislike for "foreigners," especially those of a nationality different from that of the members of the union. It is clear, however, that the predominant motive was not prejudice but economics.

Provisions against Nazis, Fascists, and certain other groups seem to be based almost wholly on the social attitudes of the members. These groups were no particular economic threat, nor is there any showing that they constituted any special danger to union organization. Such provisions might well be classified as political tub-thumping.

Denial of the right to join has been severely censured as being unjustified and undemocratic. It is not so clear on reflection, however, that unions should be condemned for refusing to admit new members when many of their present members are unemployed. Likewise, it is impossible to expect unions to freely admit those who it believes will destroy the solidarity and cohesiveness which are essential for effective action; nor should we be shocked at seeing our undesirable social attitudes reflected in union policies.

This does not imply approval of the union's denying any person the right to join; it is simply an attempt to make clear the reasons for denial. Any sound solution to the problem must be based upon an understanding of the sources from which exclusionary policies spring.

7. Northrup, op, cit., p. 5. This is especially true in the railroad unions which were founded as beneficial societies. Admission to these would have indicated social equality as contrasted with economic equality. See also Wolfe, op. cit., p. 121.

8. See generally Cayton and Mitchell, op. cit., Northrup, op. cit., and Weaver, "Recent Trends in Negro Union Relationships," Journal of Political Economy, Vol. 52, p. 234.

9. Franklin, op. cit., pp. 315, 316.

Economic Limitations on Exclusionary Policies

If a union excludes a large number of qualified workmen, thus putting them beyond all union protection and control, it may thereby defeat its own ends. These non-union workers constitute a danger to the union in two respects. First, they will inevitably seek work in non-union shops, where because they lack the bargaining power of the union, they may be compelled to work for lower wages. This may force the union to reduce the union scale in order to enable unionized shops to compete. If the union has contracts with all competitors in the industry, this danger may be temporarily avoided, but the presence of cheaper non-union labor invites the entry of new competitors. Thus aliens[1] and women[2] were not only admitted but energetically organized by the unions in the garment trades, in order to prevent their undermining union standards by working in non-union shops at low wages.

Second, whenever such a union attempts to bargain with the employer much of its bargaining power is lost, for in case of a strike the excluded workers constitute a ready supply of strikebreakers. They see it as an opportunity to better themselves, and their exclusion from the union prevents them from having any sense of solidarity with the strikers.[3] Thus Negroes were admitted to the steel and other industrial unions only after strikes were lost through the use of Negro strikebreakers.[4]

It is clear that exclusionary policies are practiced in many situations which this economic logic fails to explain. The union may be aided by employer coöperation. In many cases where the closed union has been grossly abused, it has been encouraged by employers, who have used it as a monopoly device to exploit for their own ends.[5] Thus the manufacturers and contractors com-

1. Wolfe, op. cit., p. 99.
2. Lorwin, op. cit., pp. 108, 305. In 1924 the Barbers had to amend their rules and admit women because of their influx into the trade. Wolfson, op. cit., p. 66.
3. National unions have opposed the use of permit cards partly because of the danger that the permit men would be used by the employers as strikebreakers. Slichter, op. cit., p. 69.
4. Northrup, op. cit., p. 176; Cayton and Mitchell, op. cit., passim. Unions have stated that admission was a recognition that Negroes are "economic equals," but they have denied that admission in any way involves treating them as "social equals." Auxiliary unions are an attempt to control the Negro worker without admitting him, thus protecting white workers in their job preferences.
5. Toner, op. cit., p. 80. This was particularly true in the building trades in Chicago, New York, and San Francisco.

bined with Local No. 3 of the IBEW to obtain a complete monopoly on all electrical work in New York City.[6] In a few cases unions have extended their power beyond the normal economic limitations by the use of violence against any who dared to challenge their monopoly. Assaults, bombings, and murder have been used by racketeering labor leaders to maintain their control.[7]

This presentation of economic limitations is based upon the assumption that a single union in a single industry stands alone, but this is not always the case. A single international union may cover many related industries, and may be affiliated with other unions which cover a wide range of industries. Their combined strength when mobilized may be overwhelming.[8] Restrictive practices in the building trades have been successful, not so much because of the ability of any one union to control the supply of labor as because the members of the various unions refused to use any material or work on any job which was "tainted" with nonunion labor. Few employers found it profitable to attempt to operate on a totally non-union basis. If only a small fraction of the industry is organized, the potential competition of non-union labor limits the union's ability to exclude, but this does not necessarily follow when unions become as strong as they are today. If the Auto ·Workers refused to work on non-union steel, and the Steel Workers refused to work with non-union coal, and all union members refused to buy non-union cars, few employers in these industries would risk operating on a non-union basis, even though many non-union workers were available.

Economic forces have in the past tended to limit the ability of a union to deny the right to join, but those forces are probably not now so effective as they have been. It is possible, therefore, that restrictive practices will increase. They almost certainly will increase if unemployment becomes sufficiently widespread for workers to feel a strong need for job protection.[9]

6. See Allen Bradley Co. v. Local Union No. 3, 325 U. S. 797, 65 S. Ct. 1533 (1945).

7. Seidman, op. cit., Chap. XVI. In many of these cases there was also active employer coöperation.

8. For a vivid example of the effectiveness of united action by various unions, see Auburn Draying Co. v. Wardell, 227 N. Y. 1, 124 N. E. 97 (1919).

9. The increase of restrictive membership policies has been predicted by Tipton, "Closed Unions in Modern America," Social Science, Vol. 17, p. 81, and American Civil Liberties Union, op. cit., p. 20.

Conclusions

No attempt to make a quantitative analysis of the admission policies of unions can arrive at more than approximate results. Union constitutions can be analyzed with some definiteness, although the vagueness of many of the provisions makes it difficult to determine their meaning with any certainty; but union practices are impossible to discover with accuracy, for this is not the kind of information which union officials will report objectively, even if they know the practices of all member locals. However, there are enough fairly accurate data to enable us to formulate some generalizations which are reasonably reliable.

Exclusionary policies are seldom found in industrial unions, but are restricted mainly to the skilled craft and railroad unions. Exclusion is far more prevalent in the older and well-entrenched unions than in the newer and less secure unions. In accord with both of these generalizations is the fact that provisions and practices denying the right to join are found almost exclusively in the AF of L and independent unions, and are only rarely found in the CIO unions.

Of the groups excluded, Negroes are unquestionably discriminated against most severely. Exclusion of aliens and women is of secondary importance, while exclusion on political and religious grounds is almost negligible. The extent to which unions completely close their membership is unknown, but it is probably of equal or greater importance than exclusion because of race.

It is clear that only a small minority of unions engage in exclusionary practices of any kind. The great majority of unions freely admit all workers within their jurisdiction who desire to join. That denial of the right to join is not typical of the labor movement is probably due as much to the economic factors which limit the ability of unions to exclude as it is to any political idealism. However, as unions grow stronger they will become more able to effectively limit their membership. The principal reason for exclusionary policies is to gain for the union members a degree of job and wage security in an economic system which produces mass unemployment and numerous low-wage industries. So long as these underlying reasons exist, there is considerable danger that unions will exercise their increased ability to deny admission in order to protect their older members.

Clyde W. Summers.

University of Toledo

TABLE I

ADMISSION PROVISIONS IN UNION CONSTITUTIONS[1]

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Fees Minimum	IX Initiation Fees Fixed	IX Initiation Fees Maximum
1. Actors and Artists of America, The Associated* (AF of L)	16,000	a	.....	.....	.....	.....	.....	.....	.....	.....	.....
2. Air Line Dispatchers Association (AF of L)	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
3. Air Line Mechanics Association, International* (Ind.)	45,000*	x	x	x	x	a	.....	.....	.....	$30	.....
4. Air Line Pilots Association, International* (AF of L)	2,700	.....	.....	.....	.....	.....	.....	.....	.....	10	.....
5. Aluminum Workers Union, International Council of* (AF of L)	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
6. Aluminum Workers of America* (CIO)	55,000*	.....	.....	.....	e	.....	.....	a(1)	.....	100	.....
7. Architects and Draftsmen's Unions, International Federation of Technical Engineers* (AF of L)	7,100	.....	.....	.....	.....	.....	.....	.....	5	.....	.....
8. Architects, Engineers, Chemists and Technicians, International Federation of* (CIO)	8,500	.....	x	x	x	x	.....	.....	$2	.....	$15
9. Asbestos Workers, International Association of Heat and Frost Insulators and (AF of L)	4,000	x	x	x	x	.....	.....	.....	2	.....	10
10. Associated Unions of America (Ind.)	5,000*	... b	.....	.....	.....	.....	a	r	25	.....	100
11. Automobile, Aircraft, and Agricultural Implement Workers of America (CIO)	1,200,000	x	x	x	e,x	.....	.....	.....	2	.....	15

No.	Union	Members										
12.	Automobile Workers of America (AF of L)	50,700				e				2		25
13.	Bakery and Confectionery Workers, International Union of America (AF of L)	89,700			b					3		25
14.	Barber and Beauty Culturists Union of America (CIO)	5,500	x	x					r	3		
15.	Barbers, Hairdressers, and Cosmetologists, International Union of America, Journeymen (AF of L)	50,000			b	e			a(1)	10		
16.	Blacksmiths, Drop Forgers, and Helpers, International Brotherhood of (AF of L)	10,000	b			d			r	5		
17.	Bill Posters, Billers and Distributors of America, International Alliance of* (AF of L)	1,600				f		a		15		
18.	Boilermakers, Iron Shipbuilders and Helpers of America, International Brotherhood (AF of L)	336,900	b		e				r	10		
19.	Bookbinders, International Brotherhood of (AF of L)	28,900			a	a			r	10	2	
20.	Boot and Shoe Workers Union* (AF of L)	40,000										
21.	Brewery, Flour, Cereal and Soft Drink Workers of America, International Union of United (AF of L)	67,457*			b,c							25
22.	Brick and Clay Workers of America, The United (AF of L)	10,000			b							1.50
23.	Bricklayers, Masons and Plasterers International Union of America (AF of L)	65,000			b			a,b	r	25		
24.	Bridge and Structural Iron Workers, International Association* of (AF of L)	105,600				a,d			r		100	100
25.	Broom and Whisk Makers Union, International* (AF of L)	300	x	x	a,y				a(2)	5		
26.	Building Service Employees International Union (AF of L)	70,000			b,c							

¹ See Explanatory Notes at end of Table.

TABLE I (Continued)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Fees Minimum	IX Initiation Fees Fixed	IX Initiation Fees Maximum
27. Cannery, Agricultural, Packing and Allied Workers of America, United (CIO)	50,500	x	x	x	x	x	x		$2		
28. Carpenters and Joiners of America, United Brotherhood of (AF of L)	600,000			b			b	r	10		
29. Cement, Lime and Gypsum Workers International Union (AF of L)	18,000	x	x	x	x	x			5		$25
30. Chemical Workers Union, International (AF of L)	46,000				a,b,c,f	y			2		15
31. Chorus Equity Association		x		x						$25	
32. Cigar Makers International Union (AF of L)	10,000	x								3	
33. Cleaning and Dye House Workers, International Association of* (AF of L)	17,100				e						
34. Clothing Workers, Amalgamated (CIO)	275,000	x	x	x							10
35. Communications Association, American (CIO)	18,000				x	x				20	
36. Coopers International Union of North America (AF of L)	5,000					y		r	5		
37. Diamond Workers Protective Union of America* (AF of L)				b				r			50
38. Die Sinkers Conference, International* (Ind.)	600							r	5		

	Membership											
39. Distillery, Rectifying and Wine Workers International Union (AF of L)	10,000	⋮	⋮	⋮	a	⋮	⋮	⋮	⋮	2	⋮	⋮
40. Dyers, Finishers, Printers and Bleachers of America, Federation of (CIO)	35,000*	⋮	⋮	⋮	f	⋮	⋮	⋮	⋮	⋮	⋮	⋮
41. Editorial Association, The American (AF of L)		⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1	1
42. Electrical, Radio and Machine Workers of America, United (CIO)	550,000	x	x	x	x	x	x	⋮	⋮	2	⋮	⋮
43. Electrical Workers of America, International Brotherhood of (AF of L)	312,900	⋮	⋮	⋮	⋮	⋮	⋮	b	⋮	⋮	⋮	⋮
44. Elevator Constructors, International Union of (AF of L)	10,200	⋮	⋮	⋮	⋮	⋮	r	⋮	⋮	⋮	200	⋮
45. Engineers, International Union of Operating (AF of L)	100,000	⋮	⋮	⋮	⋮	a	⋮	⋮	⋮	5	⋮	a
46. Engravers and Sketchmakers, Friendly Society of* (Ind.)		⋮	⋮	⋮	⋮	⋮	r	⋮	⋮	⋮	35	⋮
47. Farm Equipment and Metal Workers of America, United (CIO)	72,000	x	x	x	x	x	⋮	⋮	⋮	2	⋮	15
48. Farm Labor Union, National (Ind.)	20,000*	x	x	x	x	y	⋮	⋮	⋮	1	5	10
49. Federal Employees, National Federation of (Ind.)	75,000*	⋮	⋮	⋮	f	y	⋮	⋮	⋮	1	⋮	1
50. Federal Workers of America, United* (CIO)	25,000	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
51. Fire Fighters, International Association of (AF of L)	40,500	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	1
52. Firemen and Oilers, International Union of* (AF of L)	52,700	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	3.50	⋮	⋮
53. Fishermen and Allied Workers of America (CIO)	10,000	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	10
54. Foremen's Association of America* (Ind.)	15,000*	x	x	x	x	x	⋮	⋮	⋮	1	⋮	15
55. Fur and Leather Workers Union, International	90,000	x	x	x	x	x	⋮	⋮	⋮	⋮	⋮	⋮

TABLE I (Continued)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Fees Minimum	IX Initiation Fees Fixed	IX Initiation Fees Maximum
56. Furniture Workers of America, United* (CIO)	45,000	x	x	x	x	x		r	$2		
57. Garment Workers of America, United* (AF of L)	40,000						a		2		$15
58. Garment Workers Union, International Ladies (AF of L)	237,500					y					
59. Gas, Coke, and Chemical Workers of America, United (CIO)	30,000	x	x	x							
60. Glass Bottle Blowers Association (AF of L)	24,000				a,b,c,f				2		5
61. Glass, Ceramic and Silica Sand Workers of America, Federation of (CIO)	30,000			x					2		25
62. Glass Cutters League of America, Window (AF of L)	1,600							a(3)			
63. Glass Workers Union of North America, American Flint* (AF of L)	25,600			f				c		$25	
64. Glove Workers Union of North America, International* (AF of L)	3,100			a						3	
65. Government Employees, American Federation of (AF of L)	26,800			a					1		
66. Grain Processors Council, American Federation* (AF of L)	......								2		

Union	Members	1	2	3	4	5	6	7	8	9
67. Granite Cutters, International Association of America (AF of L)	4,000						r		25	
68. Hatters, Cap and Millinery Workers International Union, United (AF of L)	32,000	x	x		y					b
69. Hodcarriers, Building and Common Laborers Union of America, International (AF of L)	333,100							5		50
70. Horseshoers of United States and Canada, International Union of Journeymen* (AF of L)	200		b				r	50		
71. Hosiery Workers, American Federation of (CIO)	40,000*	x	b			b	r			
72. Hotel and Restaurant Employees International Alliance and Bar Tenders International League of America (AF of L)	224,500	x	x	e	x	a		2	c	25
73. Industrial Trades Union of America (Ind.)	14,800*	x		x	x					
74. Inland Boatmen's Union of the Pacific (CIO)	3,000	x	x					10		
75. Jewelry Workers Union, International* (AF of L)	6,900	x	x	x			r	1.50		75
76. Lace Operatives of America, Amalgamated* (Ind.)	2,875*						r			
77. Lathers, International Union of Wood, Wire and Metal (AF of L)	8,100	x	b					1		
78. Laundry Workers International Union* (AF of L)	50,000					a	a(2)	2.50		100
79. Leather Workers International Union, United* (AF of L)	2,800									
80. Letter Carriers, National Association of (AF of L)	60,000							2		25
81. Letter Carriers, National Federation of Rural* (AF of L)	400									
82. Letter Carriers Association, National Rural (Ind.)	28,066*		a							
83. Licensed Officers of the U.S.A., United* (Ind.)	2,000*								d	
84. Lithographers of America, Amalgamated*(AF of L)	13,700						r		10	
85. Locomotive Engineers, Grand International Brotherhood of* (Ind.)	76,000	a							5	

TABLE I (*Continued*)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I — Number of Members	II — Race	III — Creed	IV — Citizenship	V — Political	VI — Sex	VII — Competency	VIII — Apprenticeship	IX — Initiation Fees Minimum	IX — Initiation Fees Fixed	IX — Initiation Fees Maximum
86. Locomotive Firemen and Enginemen, Brotherhood of (Ind.)	119,686	a,c	....	....	....	....	....	....	....	$5	....
87. Longshoremen's Association, International* (AF of L)	61,000	....	....	b	....	....	....	....	$10	....	....
88. Longshoremen's and Warehousemen's Union, International (CIO)	50,000	x	x	x	x	y	....	....	....	....	$10
89. Luggage, Belt and Novelty Workers Union, International Ladies Handbag (AF of L)	15,000	....	....	....	....	....	....	....	....	....	....
90. Machine Printers Beneficial Association (Ind.)	1,400*	....	....	....	....	....	....	r	1	25	....
91. Machinists, International Association of (Ind.)	665,900	d	....	....	....	y	....	r	5	....	....
92. Maintenance of Way Employees, Brotherhood of (AF of L)	116,900	b	....	....	....	....	....	....	....	3	....
93. Marble, Slate and Stone Polishers, Rubbers and Sawyers, Tile and Marble Setters' Helpers, and Terrazzo Helpers, International Union of* (AF of L)	4,500	....	....	b	a	....	....	r	....	....	....
94. Marine Cooks and Stewards, National Union of (CIO)	6,000	x	x	x	x	y	....	....	....	32	....

	Union	Membership												
95.	Marine Engineers Beneficial Association, National* (CIO)	11,000									b		50	
96.	Marine Firemen, Oilers, Watertenders, and Wipers Association, Pacific Coast (Ind.)	3,000*				a	b						50	
97.	Marine and Shipbuilding Workers of America, Industrial Union of (CIO)	400,000	x	x	x	x	x	x				2		10
98.	Maritime Union of America, National* (CIO)	85,000	x	x	x	x	x	x					e	
99.	Masters, Mates and Pilots of America, National Organization of (AF of L)	3,000			a	a		e				25		
100.	Master Mechanics and Foremen of Navy Yards and Naval Stations, National Association of* (AF of L)	300				a	e							
101.	Meat Cutters and Butcher Workmen of North America, Amalgamated (AF of L)	100,500										5		
102.	Mechanics Educational Society of America* (Ind.) (b)	42,000*	x											
103.	Messengers, The National Association of Special Delivery* (AF of L)	900								b				
104.	Metal and Enamelware Workers, International Council of Fabricated* (AF of L)								r	b	r			25
105.	Metal Engravers Union, International* (AF of L)	900*	x						r					
106.	Metal Polishers, Buffers, Platers and Helpers International Union (AF of L)	10,000							r	b	r	5		
107.	Mine, Mill and Smelter Workers, International Union of* (CIO)	150,000	x	x	x	x	g				a(4)			
108.	Mine Workers of America, Progressive (AF of L)	35,000	x	x	x	x	b			b			50	
109.	Mine Workers of America, United (AF of L)	600,000	x	x	x	g							10	
110.	Molders and Foundry Workers Union of North America, International (AF of L)	65,700	x		x							5		

TABLE I (*Continued*)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Feed Minimum	IX Initiation Feed Fixed	IX Initiation Feed Maximum
111. Musicians, American Federation of (AF of L)....	100,000	..	..	b,c	a,b,c	y	..	..	$1	..	$50
112. Newspaper Guild, American (CIO)............	21,000	x	x	x	x	x	..	..	..	..	10
113. Office Employees International Council, American Federation of (AF of L)..................	....	..	..	..	..	..	..	..	..	..	..
114. Office and Professional Workers, United* (CIO)....	55,000	x	x	..	x	..	..	..	..	f	..
115. Oil Workers International Union (CIO)........	53,000	x	x	..	a,c	x	..	..	2	..	25
116. Packinghouse Workers of America, United (CIO).	92,000	x	x	x	y	x	x	..	..	$2	..
117. Painters, Decorators and Paper Hangers of America, Brotherhood of (AF of L).............	139,700	..	..	..	a,b,c	..	b	a(3)	5	..	..
118. Paper Makers, International Brotherhood of (AF of L).................................	34,400	..	..	..	..	..	b	..	8	..	50
119. Paper, Novelty and Toy Workers International Union, United* (CIO)...................	12,000	..	..	x	..	..	..	..	..	..	10
120. Paper Workers Organizing Committee (CIO)	12,000	x	x	..	..	..	..	..	..	2	..
121. Pattern Makers League of North America* (AF of L).................................	11,000	..	..	..	..	..	a	a(5)	5	..	..
122. Paving Cutters Union of United States and Canada* (Ind.)............................	1,800*	..	..	f	..	..	..	a(2)	..	10	..
123. Photo-Engravers Union of North America (AF of L)	10,900	..	..	..	..	y	..	a(6)	..	..	..

No.	Organization	Membership	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10
124.	Plasterers and Cement Finishers, International Association Operative (AF of L)............	25,000	.....	.....	.....	.....	.....	a	.....	.....	.....	g
125.	Plate Printers, Die Stampers and Engravers Union of North America, International* (AF of L)....	1,100	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
126.	Plumbers and Steamfitters of United States and Canada, United Association of (AF of L)......	130,000	.....	.....	.....	.....	.....	a	a(4)	20	.....	.....
127.	Postmasters, National League of District (Ind.)....		.....	.....	.....	.....	.....	.....	a(4)	.....	.....	.....
128.	Post Office Clerks, National Federation of* (AF of L)............	40,000	.....	.....	.....	.....	.....	.....	a(5)	.....	.....	.....
129.	Post Office Clerks, United National Association of (Ind.)............		.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
130.	Post Office and Railway Mail Laborers, National Association of* (AF of L)............	1,500	.....	.....	.....	e	.....	.....	.....	.....	.....	.....
131.	Postal Employees, National Alliance of (Ind.)....	10,047*	.....	.....	.....	.....	.....	.....	.....	.....	0.65	.....
132.	Postal Supervisors, National Association of (Ind.)	9,400*	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
133.	Potters, National Brotherhood of Operative* (AF of L)............	21,500	.....	.....	.....	.....	.....	.....	.....	.....	.....	.....
134.	Printing Pressmen and Assistants Union of North America, International* (AF of L)............	49,300	.....	.....	.....	.....	.....	c	r	.....	3	.....
135.	Pulp, Sulphite and Paper Mill Workers, International Brotherhood of (AF of L)............	61,300	.....	.....	.....	.....	.....	.....	.....	2	.....	15
136.	Railroad Signalmen, Brotherhood of (Ind.)......	13,000*	.....	b	.....	.....	.....	.....	.....	.....	.....	.....
137.	Railroad Telegraphers, Order of (AF of L)........	30,000	a	.....	.....	.....	.....	.....	.....	5	5	.....
138.	Railroad Trainmen, Brotherhood of (Ind.)........	210,570	a	.....	.....	.....	a	.....	.....	1	.....	.....
139.	Railroad Yardmasters of America* (Ind.)........	3,500	.....	.....	.....	.....	a	.....	.....	.....	10	.....
140.	Railway Carmen of America, Brotherhood of* (AF of L)............	95,800	a,b	.....	.....	.....	.....	.....	r	.....	.....	10

TABLE I (*Continued*)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Fees Minimum	IX Initiation Fees Fixed	IX Initiation Fees Maximum
141. Railway and Steamship Clerks, Freight Handlers, Express Station Employees, Brotherhood of (AF of L)	204,200	a	....	....	....	y	....	....	....	$5	....
142. Railway Conductors of America, Order of (Ind.)	36,000	a	....	....	....	a	....	....	....	....	$5
143. Railway Mail Association (AF of L)	21,800	e	....	....	....	a	....	....	....	4.25	....
144. Railway Patrolmen's Union, National Council* (AF of L)	....	....	....	....	....	....	....	....	....	....	....
145. Railway Supervisors Association, Inc., The American (Ind.)	2,500	....	....	....	....	....	....	....	$3	....	....
146. Retail Clerks, International Protective Association (AF of L)	100,000	....	....	....	....	....	....	....	3	....	....
147. Retail, Wholesale and Department Store Employees of America, United (CIO)	120,000	x	x	....	x	x	....	....	....	....	....
148. Roofers, Damp and Waterproof Workers Association, United Slate, Tile and Composition* (AF of L)	8,000	....	....	b	e	....	....	....	25	....	10
149. Rubber, Cork, Linoleum and Plastic Workers of America, United (CIO)	200,000	....*	....	....	a,b,c,e	....	....	....	....	2	....

No.	Union	Membership	1	2	3	4	5	6	7	8	9	10
150.	Seafarers International Union of North America (AF of L)	30,000										
151.	Sheet Metal Workers, International Association* (AF of L)	25,000	b		b			b	r	25		h
152.	Shoe and Allied Craftsmen, Brotherhood of (Ind.)	11,000				x	y			2	2	
153.	Shoe Workers of America, United* (CIO)	65,000	x	x	x					5		
154.	Siderographers, International Association of* (AF of L)	100					x		a(5)		15	
155.	Sleeping Car Porters, Brotherhood of* (AF of L)	9,000	x	x	x							
156.	Spinners Union, International* (AF of L)	500										
157.	Stage Employees and Moving Picture Machine Operators of the United States and Canada, International Alliance of Theatrical* (AF of L)	42,000				e		a	r			j
158.	State, County and Municipal Employees, American Federation of (AF of L)	46,000	x	x			x, y	a		2		10
159.	State, County and Municipal Workers of America* (CIO)	44,500	x	x	x	x				1		5
160.	Steel Workers of America, United (CIO)	1,000,000	x	x	x		x	a			3	
161.	Stereotypers and Electrotypers Union of North America, International (AF of L)	8,800			b				a(6)			
162.	Stone and Allied Products Workers of America, United (CIO)	6,000	x	x	x	x				5		
163.	Stone Cutters Association of North America, Journeymen* (AF of L)	1,900			b,g				a(4)		75	
164.	Stove Mounters International Union of North America (AF of L)	8,000							a(3)	3.50		25

TABLE I (Continued)

ADMISSION PROVISIONS IN UNION CONSTITUTIONS

Name of International Union	I Number of Members	II Race	III Creed	IV Citizenship	V Political	VI Sex	VII Competency	VIII Apprenticeship	IX Initiation Fees Minimum	IX Fixed	IX Maximum
165. Street, Electric Railway and Motor Coach Employees of America, Amalgamated Association of* (AF of L)	104,800	a	...	...	...	...	...	...	$2	...	...
166. Switchmakers Union of North America* (AF of L)	9,300	a	...	...	...	a	...	...	2	...	...
167. Teachers, American Federation of (AF of L)	25,200	x	x	...	a,b,c,x	...	...	...	...	...	...
168. Teamsters, Chauffeurs, Warehousemen and Helpers of America, International Brotherhood (AF of L)	629,200	...	...	b	a,f	...	...	...	...	...	...
169. Telegraphers Union of North America, Commercial (AF of L)	18,700	...	...	...	a,b,c	...	...	...	5	...	...
170. Telephone Workers, National Federation of (Ind.)	167,000*	...	...	...	...	...	...	...	...	...	...
171. Textile Workers of America, United (AF of L)	37,200	...	...	...	...	y	...	...	...	...	...
172. Textile Workers Union of America (CIO)	325,000	...	...	...	...	...	...	...	...	$1	...
173. Tobacco Workers International Union* (AF of L)	22,000	...	...	...	...	y	...	...	1	...	...
174. Tool and Die Craftsmen, Society * (Ind.)	...	...	...	...	...	...	...	...	5	...	...
175. Train Dispatchers, American Association of (Ind.)	3,500*	...	...	...	...	...	...	...	...	10	...
176. Transport Service Employees of America, United (CIO)	15,000	x	x	x	x	...	...	...	...	5	...
177. Transport Workers Union of America (CIO)	95,000	x	x	x	x	x	...	...	3	...	...

178. Typographical Union, International (Ind.)......	83,500*	...	...	...	...	...	...	a	a(6)	...	35	...
179. Upholsterers International Union of North America* (AF of L)......	25,000	...	...	...	...	y	...	...	2	...	2	$25
180. Utility Workers Organizing Committee* (CIO)...	35,000	x	x	...	...	y	...	...	...	...	...	...
181. Wall Paper Craftsmen and Workers of North America, United (AF of L)......	2,800	x	x	b,h,x	...	y	...	...	...	3	...	100
182. Welders of America, United Aircraft* (Ind.)......	3,000*	x	x	a	a,b,c, d,e	...	...	...	...	...	...	15
183. Wire Weavers Protective Association, American (AF of L)......	400	...	...	b,d	...	a	...	a(4)	...	...	25	...
184. Wood Carvers Association of North America* (AF of L)......	300*	x	x	...	...	y	...	r	25	...	...	...
185. Woodworkers of America, International (CIO) ...	60,000	x	x	x	a,b,c	y	...	...	5	...	...	...

TABLE I

EXPLANATORY NOTES

This analysis is taken directly from the international constitutions where they were available. Where no constitution was available (marked by an asterisk after the name of the union), the analysis was taken from Peterson, Handbook of Labor Unions (1944). The analysis does not purport to reflect union practices, but only the constitutional provisions. If no provision is indicated, the local unions are free to establish such rules and regulations as they see fit. The number of each kind of provision is indicated below in parenthesis after each provision. The membership listed in Column I is taken from Labor Fact Book, No. 7 (1945) where available. Those marked with an asterisk were taken from Peterson. The total membership of the unions listed is 13,566,701.

Col. II. Provisions Concerning Race.
     (a) "white" (9)
     (b) auxiliary locals for Negroes (5)
     (c) "Mexicans, Indians, or those of Indian or Spanish-Mexican extraction excluded" (1)
     (d) Negroes excluded by ritual (1)
     (e) "Caucasian race or native American Indian" (1)
     (x) "eligible regardless of race," "irrespective of race," "not barred because of race," "no person is to be barred because of race," and similar provisions which prohibit exclusion (47)

Col. III. Provisions Concerning Creed
     (a) "Firm believer in God, the Creator of the Universe" (1)
     (b) "believes in the existence of a Supreme Being" (1)
     (x) "eligible regardless of creed," etc. (see provisions concerning race) (46)

Col. IV. Provisions Concerning Citizenship and Nationality
     (a) citizenship required (7)
     (b) declaration of intent or first papers required (22)
     (c) become a citizen as soon as possible required (4)
     (d) foreign wire weavers must pay an initiation fee of $1,000 (1)
     (e) "citizen of some civilized country" (1)
     (f) if trade is learned in a foreign country, must pay an initiation fee of $200 (2)
     (g) applicant from foreign country without a clear union card must pay an initiation fee of $100 (1)
     (h) if trade is learned in a foreign country, must pay an initiation fee of $300 (1)
     (x) "eligible regardless of nationality," etc. (see provisions concerning race) (2)
     (y) no person barred because of nationality "except Asiatic Labor" (1)

Col. V. Provisions Concerning Political Affiliation or Belief
     (a) member of the Communist Party or a communist organization excluded (15)
     (b) member of the Nazi Party or bund excluded (9)
     (c) member of the Fascist Party excluded (10)

  (d) member of the I.W.W. excluded (3)
  (e) member of any organization "whose aims, principles and philosophy are contrary to those of the Constitution"; "which advocates the overthrow of the Government by force"; "hostile to the American form of government"; "that expounds principles inimical to our constitutional government"; or "subversive group seeking to advance the interests of totalitarianism"; excluded (12)
  (f) excluded although not a member of any of the above groups, but for advocating such causes or purposes (6)
  (g) "members of the I.W.W., Working Class Union, One Big Union, Chamber of Commerce, Ku Klux Klan, Communist, Nazi, or Facist parties" excluded (1)
  (x) "eligible regardless of political affiliations or beliefs," etc. (see provisions concerning race) (28)
  (y) "eligible regardless of political affiliations or beliefs consistent with the democratic process" (1)

Col. VI. Provisions Concerning Sex
  (a) "male" (8)
  (x) "eligible regardless of sex," etc. (see provisions concerning race) (19)
  (y) "all men and women," "male and female" (19)

Col. VII. Provisions Concerning Competency
  (a) competency test required (11)
  (b) "able to command the union wage," or "qualified to work at the trade" (8)
  (c) "qualified to work at the trade" (1)
  (x) "eligible regardless of skill" (2)

Col. VIII. Provisions Concerning Apprenticeship or Working Experience
  Number of years required indicated in parenthesis
  (a) apprenticeship or working experience required (19)
  (b) "experienced workman" (1)
  (c) initiation fee of $125 if no apprenticeship is served (1)
  (r) constitutional regulation of apprenticeship but no requirement as to admission (30)

Col. IX. Initiation Fees
  The minimum listed is the lowest for regular membership; the maximum listed is the highest for regular membership; and the fixed listed is the highest fixed fee for regular membership.
  (a) maximum set by the General President (1)
  (b) maximum is the prevailing weekly wage of the branch for work done by the applicant (1)
  (c) set by the General Meeting (1)
  (d) set by the Executive Committee (1)
  (e) set by the National (1)
  (f) set by the local but approved by the Executive Board (1)
  (g) maximum of 10 times the daily wage (1)
  (h) maximum of 100 hours of labor at minimum wages (1)
  (j) maximum of 4 times the highest regular weekly wage in the local's jurisdiction (1)

MONOPOLY PRICING AND UNEMPLOYMENT

SUMMARY

Introduction: neglect of monopoly influence, 108. — The model, 109. — Initial assumptions, 110.— Primary unemployment, 110.— Rising cost curves, 112. — Variations in aggregate consumer outlay, 114. — Secondary unemployment, 117. — Derived demand for capital goods, 120. — Monopoly in investment goods, 121.— Reconciliation with Keynes' theory, 121. — Conclusion, 124.— Appendix of symbols, 124.

INTRODUCTION

Monopoly is receiving an increasing amount of criticism as a prime cause of unemployment.[1] Yet, in seeking the forces determining the level of activity, studies in the theory of employment have, by and large, subordinated the special influence of monopoly. Lord Keynes subsumed the degree of competition as part of the data, along with consumer tastes, productive technique and the quantity and quality of labor and equipment.[2] Mrs. Robinson, presenting merely the major elements of Keynesian theory, makes only passing reference to the rôle of monopoly.[3] Even the recent work of Dr. Lange deals mainly with the question of the rigidity of factor prices as an unemployment force, not with monopoly prices of finished goods.[4] Mr. Harrod, on the other hand, imputes a primary rôle to monopoly prices in determining the movement of prices and profits consequent upon a change in the level of economic activity.[5] However, some of his main contentions rest upon the assumption of a diminishing elasticity of demand as demand curves move to the right in the upward phase of the business cycle; and the à priori validity of this proposition has been subject to criticism ever since Mr. Harrod first stated it.[6]

1. As in President Truman's budget message, New York Times, January 22, 1946.
2. J. M. Keynes, The General Theory of Employment, Interest and Money (New York, 1936), p. 245.
3. Joan Robinson, Introduction to the Theory of Employment (London, 1937), pp. 51–53.
4. Oscar Lange, Price Flexibility and Employment (Bloomington, Indiana, 1944), p. 1.
5. R. F. Harrod, The Trade Cycle (London, 1936), p. 85.
6. E.g., R. F. Bretherton, "A note on the Law of Diminishing Elasticity of Demand," Economic Journal (1937), Vol. XLVII, pp. 574–577; H. W. Singer, "The Law of Diminishing Elasticity of Demand," ibid. (1938), Vol.

It is quite possible, however, to analyze the relation of monopoly pricing to employment without making use of Harrod's postulate. It is, for example, interesting and important to inquire as to the effect of monopoly upon employment at a given level of investment activity, rather than posit its stabilizing effects, as Harrod has done, at changing income or investment levels. Normally, this will be a fundamental characteristic of the model which we shall investigate, a contrast and supplement to Mr. Harrod's procedure.

THE MODEL

We shall start from the individual firm, which has come to be a bedrock for analysis,[7] and then trace the consequences for total income and employment when, after previously equating marginal costs (MC) to price (P), the firm decides to exercise the monopoly power it possesses and equates MC to marginal revenue (MR), where $P > MR$. After examining the effects upon employment of a single firm initiating monopoly pricing practices, added weight can be attributed to our conclusions by supposing that n firms behave in the same way simultaneously.

The advantage of assuming that the firms are initially in an optimum price position rests on our familiarity with this model. After introducing monopoly pricing into the system, and analyzing its effects on the general equilibrium, the consequences of a reverse movement from a monopoly to a more competitive type of behavior ought to be readily discernible. There is no gainsaying that the latter implications are the relevant ones for social policy. Our more comprehensive analysis, we shall find, will largely confirm impressions of the effects of monopoly upon employment based upon simpler single-firm equilibrium, although the degree of unemployment is never as definite and clear-cut as the latter would lead us to believe, being sometimes somewhat more, and frequently less. Occasionally, rather than involve unemployment, there are instances in which monopoly may increase aggregate employment.

XLVIII, pp. 138–141; H. Makower, "Elasticity of Demand and Stabilization," Review of Economic Studies (1938), Vol. 6; J. D. Sumner, "Cyccilal Changes in Demand Elasticity," American Economic Review (1940), Vol. 30.

7. See Robert Triffin, Monopolistic Competition and General Equilibrium Theory (Cambridge, 1940), p. 93. Also, K. Boulding, "The Theory of the Firm in the Last Ten Years," American Economic Review (December 1942), Vol. XXXII, p. 791.

Some clues to industries in which monopoly pricing may be particularly harmful can be deduced, it is believed, from the abstract discussion to follow.

INITIAL ASSUMPTIONS

Before going to the core of the problem, the initial simplifying assumptions may be listed. The firm whose actions are under scrutiny for its departure from a $P = MC$ price policy will be referred to as A, or the A firm. That part of the economic field, or those firms, affected by A's actions will be denoted as the B firms. The rest of the economic world, which experiences no repercussions whatsoever, may be termed the C firms. The initial assumptions follow: (1) for all firms in the system, including A, marginal costs are presumed constant; further, marginal costs consist entirely of wage costs; (2) wage rates are given, fixed for each individual; (3) at the outset, to isolate their behavior from the rest of the system, we suppose that all those originally employed by A spend nothing on consumption, either before or after A's shift to a $P > MR = MC$ policy; (4) those employed in B and C make the same aggregate consumption outlay before and after A prices monopolistically; (5) the monopoly income recipients in A spend either nothing or an unchanged amount on consumption both before and after the shift to a $P > MR = MC$ price; (6) a constant volume of investment activity continues, undertaken perhaps by the C firms or by government, despite A's new price policy; (7) all the B firms are presumed to behave competitively, to equate P to MC. Although our assumptions abstract some vital elements, they simplify our problem enormously. They will be removed at easier, later stages.

PRIMARY UNEMPLOYMENT

The demand for A's product, we can posit, is elastic. As a result of the $P > MR = MC$ price now named, total expenditure on A's good falls off; aggregate expenditure being constant, the outlay on the products of the B firms will increase by exactly this amount. Employment also moves in these directions, following the flow of expenditure. More quantitative statements, however, are necessary to understand the net effects for employment. From the above statement we might draw the erroneous conclusion that total employment is unaffected.

Expenditure on A's output has decreased from $P_1 X_1$ to $P_2 X_2$,

where $P_2 > P_1$ while $X_2 < X_1$.[8] The difference, $P_1X_1 - P_2X_2 = \Delta D_B$, measures the amount of demand shifted to the B firms. But total wage incomes in A, which formerly comprised A's full income P_1X_1, now fall to P_1X_2, a decrease of $P_1(X_1 - X_2)$. P_1, it will be recalled, equals marginal costs which were previously postulated as constant and comprised solely of wage costs. The difference $P_1(X_1 - X_2)$, or the fall in the wage bill in A, exceeds ΔD_B by exactly the amount of monopoly profits, I_A equal to $X_2(P_2 - P_1)$. Dividing I_A by W, the average wage rate, we can determine the volume of *primary* unemployment induced in the system by A's monopoly price policy, on our assumptions thus far.[9] The primary unemployment is somewhat less than the loss in employment in A, or less in amount $\Delta D_B/W$.

Thus, with constant investment and constant consumer outlay, and in a world of constant MC, after the shift to monopoly total employment is smaller by the quotient $\dfrac{\text{monopoly profits}}{\text{average wage rate}}$; thus primary unemployment $U_P = I_A/W$. We conclude, therefore, that with given wage rates, if total employment is to be the same in the monopoly as in the optimal price world, total income in the former will have to exceed that of an optimal world by an amount equal to the sum of monopoly profits.

Since we are dealing with a problem of total employment, certain comments on the income level, relative savings positions, and monetary policy may not be amiss. As investment and consumer expenditure were assumed unchanged in moving from the one to the other world, total income is constant; there are thus no changes in the aggregate money supply. Total savings are likewise unchanged and equal to total investment. But as we assumed that the now unemployed group, U_P, had previously spent nothing on consumption, we likewise suppose that all of the income which now emerges as monopoly profits, I_A, represents savings. If the U_P group formerly purchased securities with their full income, we can assume I_A to be diverted to this purpose. There is thus no change

8. A's demand, unless it sells an "inferior" good, must be elastic when total consumer outlay is constant. See below (p. 114) for some remarks on situations of inelastic demand.

9. In Mrs. Robinson's terminology, primary unemployment means something different, being the direct unemployment in investment good output resulting from a decrease in investment activity. Op. cit., pp. 18–20.

in the aggregate liquidity preference and, with the constant money supply, interest rates are unchanged. If there is a change in the demand for securities, we can rule out any new effects from this direction by supposing a banking policy designed to keep interest rates constant — a not unrealistic assumption.

There is this further justification for the assumption of a constant interest rate policy, which we shall make throughout. In practice, interest rates are to a large extent amenable to banking control. If we find monopoly pricing responsible for unemployment through another channel, it is hardly an offset to suggest that, nevertheless, it should lead to lower interest rates, which in turn should stimulate employment. For any increase in employment from this direction could be secured, without monopoly pricing and its concomitant unemployment drags, through banking policy. Effects of monopoly pricing upon the interest rate structure would only be important in an economy in which the supply of money was fixed and banking policy was impotent to alter it.

RISING COST CURVES

The preceding analysis was based on the supposition that MC curves in all outputs were constant. This assumption can now be removed and the necessary modifications made in the analysis.

First, we can suppose that in A the MC curve is rising, while still constant in B. Mrs. Robinson has demonstrated that if a rising (linear) MC curve intersects a demand curve at the same point as a curve of constant MC, then the monopoly output of the firm will be larger and price lower.[1] Avoiding a discussion of any redistribution of rent incomes in A as a result of its new $MR = MC$ price policy, it is clear that the loss in employment will equal I_A minus any diminution in rents (or "non-wage" income), divided by the average wage rate. Output greater and monopoly profits smaller, primary unemployment will be reduced. If the MC curve is rising in both A and B, then prices will rise in the B firms as demand is shifted there after A's $MR = MC$ price policy. A's (linear) demand curve, therefore, will be less elastic[2] than when MC is constant in B; A's monopoly price and profits should be greater, a bad augury for employment.

1. Joan Robinson, Economics of Imperfect Competition, Chapter 2.
2. This, and the remarks to follow, can be illustrated by means of a pair of interrelated demand curves such as I have used in my "Foundations of the Demand Curve," American Economic Review (September, 1942), pp. 545–547.

The important new point is that the transfer of demand to B, ΔD_B, has raised prices in B. Part of ΔD_B, therefore, goes not to further employment in B and to offset the unemployment in A, but rather to enhance rents or short-period profits in B. Unemployment is now equal to the sum of monopoly profits in A, minus any diminution in rent incomes there (ΔR_A), plus the increased rent incomes in B, ΔR_B, divided by the average wage rate. Thus:

$$U_P = \frac{I_A - \Delta R_A + \Delta R_B}{W}.$$

Thus, when the substitutes to which expenditure is diverted after the introduction of a monopoly price policy are produced under conditions of rising marginal costs, we can expect primary unemployment in the system to be greater than when constant marginal costs prevail. Rising marginal costs dilute even the usual compensatory effects on employment of ΔD_B. On the other hand, if costs are falling in the B firms, while somehow B equates P to MC, as it could through a government subsidy to cover capital costs, this should go a long way to enforce almost an optimal price policy on A, if the commodities are fairly close substitutes; for A's demand would be rendered more elastic. Employment should also be largely maintained, for the transfer demand to B would go entirely to wage payments.[3]

We thus have a more inclusive measure of primary unemployment at the constant total income level after the advent of monopoly. As we had postulated that the U_P group formerly saved their full income, we now suppose the same savings behavior of the new income recipients, an assumption which we shall remove later. Once more, to obviate any effects on the rate of interest due to the shift in income, we assume that the aggregate demand for securities is unchanged or that banking policy irons out any tendencies to a change in interest rates. To rid this more monopolized world of the primary unemployment, total income will have to rise, so long as wage rates are constant, by the amount $I_A - \Delta R_A + \Delta R_B$; government expenditures under a compensatory fiscal policy, say,

3. When there is a vast discrepancy in the ratio of labor to other factors among the A and B firms, and when A's output requires relatively little labor, then the diversion of expenditure to B through monopoly pricing in A may expand total employment. In this case, ΔR_B will be practically zero, while $\Delta R_A > I_A$. When A is the much more labor using, ΔR_A will be practically zero, while ΔR_B will about equal ΔD_B.

will have to rise by this amount, by borrowing from the banks or receiving in taxes sums which would otherwise be "hoarded" — while the interest rate is kept unchanged through banking policy — in order to restore the original level of employment. This supposes that the Government has set itself the task of maintaining the "full" (previous) employment level.

VARIATIONS IN AGGREGATE CONSUMER OUTLAY

Until now we have supposed that aggregate consumer outlay remains constant, despite the emergence of monopoly pricing in the system. This restriction can now be removed. But to appraise the effects of thus widening our system, it will be helpful to make some arbitrary distinctions between "poor" man's goods and "rich" man's goods, although in the real world it may not be possible to distinguish goods to any great extent according to such a criterion.

First, suppose that a good taking a fairly sizable portion of the total expenditure of lower income groups is subject to monopoly pricing. Either this group fails to save at all or its savings are fixed, as through insurance and social security payments. For this case the preceding analysis based on a constant aggregate consumer outlay is particularly apposite. We have already outlined the nature of the primary unemployment that would develop in this system.

Suppose, instead, that our income group makes some limited amount of savings. In these circumstances there will probably be some diversion of income from savings to current expenditure in an endeavor to maintain consumption, especially if the monopolized item is important in the conventional standard of living. As a result, some lift will be given to total demand, and thus to employment. To ascertain the volume of unemployment attributable to monopoly pricing, after variations in aggregate consumer outlay (ΔO_T) are recognized, the addition to total outlay must be deducted from the amount of monopoly profits (and enhanced rents) before dividing by the average wage rate. Thus, with an increase in consumer outlay, and neglecting rent items, primary unemploy-

ment is: $U_P = \dfrac{I_A - \Delta O_T}{W}$.

Allowing for an increased aggregate outlay, situations of inelastic demand — the traditional monopoly plum — become

more common. For previously, with fixed outlay, unless A's commodity was an inferior good, a higher price for it shifted expenditure to B, necessarily implying elastic demand for A. But with an increased outlay, B's demand can expand even with inelastic demand in the A market. Still, no new analytical considerations are involved, at least not at this stage. It may be pointed out, however, that a portion of the increased outlay goes merely to swell monopoly profits, and thus does not by itself constitute a net gain for employment, even though, in the formula expression, it constitutes an offset to the total of monopoly profits. So long as $I_A > \Delta O_T$, the increased outlay will never entirely eradicate the primary unemployment.

This analysis suffices for expenditure changes of low income groups possessing some savings margin with which to maintain standards of living, and with respect to commodities purchased out of current income. But there is a vast area of goods, consumer durable goods, purchased only partially out of current income and mainly out of accumulated savings, such things as houses, automobiles, refrigerators, washing machines, etc. These are generally regarded as mass produced, for middle and lower income groups with limited savings margins. If monopoly pricing is prevalent in these fields, with $P > MC$, the likelihood is that demand will be elastic and total purchases will fall off. To some extent the B outputs should benefit, should experience rising demands. But to a large extent consumers will be surfeited with the B good and prefer to add to their accumulations, saving more at present and, because of the higher monopoly price, postponing the purchase of the durable good to a future period when their wealth position is better able to afford it. Immediately, however, this will mean a fall, a diminution in total consumer outlay and, compared with the previous position of a heightened or positive ΔO_T, this time it is a negative ΔO_T. Employment will now be reduced by the sum of both monopoly profits and restricted outlay, or $\dfrac{I_A + \Delta O_T}{W}$.

Thus in the field of mass produced consumer durable goods, where the alternative to current purchase is increased current savings, and where the sums involved are large, a policy of monopoly pricing will have most severe effects on employment. For here we do not find, to any great extent, a transfer of demand to

the B firms and the compensatory effects on employment already noted. Some may object that this passage presupposes cash purchases of consumer durables and overlooks the fact of installment buying. So long as the demand curve does fall to the right, however, the conclusions on unemployment are unaltered. But there is an additional point to add, for now, with installment buying, it is possible for demand to be fairly inelastic. Nevertheless, through time the heightened monopoly price will impose a greater savings burden on consumers in order to repay their larger installment loans and will reduce buying through the future of the various B commodities. Thus we have a stream of unemployment in B, through reduced outlay there, besides the restricted monopoly employment in A, so long as demand possesses some elasticity. Rather than a positive ΔD_B, we may now have a negative ΔD_B.[4]

Passing from the "poor" man's or mass produced durable goods, we can examine a class of goods consumed more exclusively by the "wealthy." For this group, we can suppose that an effort will be made to sustain consumption despite the monopolistic price policy. Cases of inelastic demand should, therefore, be more frequent and, in view of the positive ΔO_T, total employment will be better maintained. In the case of those commodities bought by this income group *solely because the price is higher* — Veblen's "conspicuous consumption" — monopoly pricing will actually *increase* total employment. Other possibilities of monopoly enlarging total employment will be uncovered by the analysis of the next section.

At this point the theorist coming to this subject of monopoly pricing from monetary or aggregate demand analysis may demur, unless certain assumptions are made very explicit. For previously, since we supposed total income or total expenditure constant, there was little need to bother with monetary policy or the rate of interest (even though we did make the underlying suppositions clear). When we permit the level of income to vary, these ideas cannot be dismissed so lightly. We have spoken of individuals spending greater amounts out of income in the more monopolized world than they did in the optimal pricing world. Thus, by necessity, their savings would be reduced, and this is likely to have an effect on the rate of interest, if they formerly purchased securities and the money supply is constant, or even if they had previously

4. I am indebted for this point to Miss Ruth Tarlow.

failed to purchase securities and held their savings idle. And although saving in amount exactly equivalent to the enhanced consumption expenditure is now done by the monopoly (and induced rent) income recipients, there will be some effects on the rate of interest, unless the aggregate liquidity preference schedule is unaltered after these movements. Once more, to dispose of any effects along these channels we suppose the banking system makes available the necessary funds to maintain the structure of interest rates. Thus when increased installment lending was called for at the higher monopoly price, we suppose the banking system to lend at an unchanged rate of interest. Or when the enhanced aggregate consumer expenditure out of incomes means a contraction in the rate of security purchases, we suppose that the banking system steps in to buy the securities which those now spending at a higher rate customarily purchased. When the increased expenditures merely mean a diminution in funds otherwise idle, and if the new monopoly income recipients are inclined to buy securities at the market rate of interest, then the banking system is presumed to sell securities (or curtail its volume of purchases) so that interest rates are kept from falling. Thus whatever the particular interactions on the security markets, monetary policy is assumed to be such as to stabilize the structure of interest rates.

Secondary Unemployment

Another limiting postulate needs to be removed, namely, the assumption that those employed by the A firm made no consumption expenditure whatever, either during A's period of $P = MC$ pricing or later during the $MR = MC$ period. Likewise, we will want to remove the proviso that none of the new monopoly income received by the A "owners" was spent. We thus concern ourselves with the further consequences on employment of the redistribution of income from the U_P group to the monopoly (and rent) income recipients.

The labor originally employed by A might be subdivided into the following categories: (1) those who remain employed by A even after the $P > MR = MC$ price policy; denoted by L_M; (2) those who find employment in B after the transfer of demand to B, denoted as ΔL_B; and (3) those who remain unemployed, or the primary unemployment U_P. A's original labor force is thus $L_C = L_M + \Delta L_B + U_P$.

So far as the L_M and ΔL_B groups are concerned, we can immediately lift our restriction and allow them to lapse into anonymity with the rest of the consuming public. They too can be assumed to expend either an unchanged amount after the new $P > MR = MC$ price policy or, like the others, to vary their consumer outlays.[5]

The U_P group, on the other hand, constitutes a harder nut. Insofar as the consumption expenditures of this group are changed (ΔC_U), lowered because of their unemployment, there will be further repercussions through the entire economic system. We will want to consider this in conjunction with the changed fortunes of the monopolists and especially any increased consumption disposition, ΔC_M, arising from their new income.[6] For both groups will probably vary their consumption outlay.

In the first place, it is obvious that the consumption expenditures of the U_P group will not fall to zero, despite the loss in employment. Either through the drain of whatever savings they possess, or through borrowing, or through public relief, some consumption expenditures continue. Roughly the same will be true of those who benefit by the monopoly pricing policy. They, too, had some income and made some expenditure prior to their monopoly foothold and the $P > MR = MC$ price innovation; they would also make some expenditures in a reverse movement, from a monopoly to a $MC = P$ pricing policy. Thus it is only the net variation in the expenditures of these groups which is important. Thus we will want to inquire whether the absolute amount $\Delta C_U \gtreqless \Delta C_M$, where ΔC_U is a decrement and ΔC_M, an increment.[7]

The case of $\Delta C_U = \Delta C_M$ may be considered first, since it prese ts fewer difficulties. Here there are no new effects on aggregate expenditure, and there is no additional unemployment beyond the primary unemployment. What has happened, however, is that real consumption formerly made by the U_P, the wage earning

---

5. They will probably figure as insignificant components of the demand for the products they help produce; otherwise, we are dealing with a case of interrelated demand and supply schedules.

6. We omit any considerations of changed rent incomes, as in the B firms as a result of A's monopoly price. The modifications in the analysis follow quite simply.

7. Merely enumerating the three possibilities indicates that the usual assertion — that monopoly causes unemployment because, for one thing, it shifts income from those with high propensities to consume to those with low — is too vague, if not positively erroneous in certain instances.

group, is now done by the monopoly income recipients. This does imply a real shift in well-being.

Case 3, of $\Delta C_U < \Delta C_M$, has largely been ruled out in customary discussions of this subject. For the monopolist is generally pictured as a wealthy individual whose current consumption bears little or no relation to current income. Nevertheless, if such a situation is found, it is clear the increased consumption outlay will tend to stimulate employment and would, through this channel, reduce the volume of (primary) unemployment attributable to the monopoly profits at the given aggregate income level. It is not really hard to envisage situations in which this analysis might be directly applicable. For example, suppose A is a monopolist catering largely to a wealthy clientele and selling, say, home furnishings. Demand here may be quite inelastic, and probably at the expense of savings. Consequently, monopoly pricing may result in little in the way of primary unemployment. On the other hand, the monopoly income may be large and shared fairly widely among all those who come under the monopolists' aegis; the owners of the business may grant income bonuses or higher wages to those employees who remain after the shift to monopoly pricing. Conceivably, consumption expenditures and employment might thereafter rise to such an extent as actually to overshadow primary unemployment.[8] Thus, in limiting cases even total employment might be expanded by monopoly pricing, just as we stated it might be with goods sought for "conspicuous consumption."

Instances in which $\Delta C_M > \Delta C_U$ should thus not be rare. In any event it is quite unrealistic, probably wrong, to treat ΔC_M as always a negligible quantity. It will depend upon the distribution of monopoly profits and the wealth position of the monopoly income recipients. In today's economy, where wage disputes involve discussions of "ability to pay," it would be surprising indeed if there were not some substantial distribution of monopoly gains among the firm's employees, for it has been in those very industries in which monopoly pricing is regarded as rife that this principle of "wage determination" has come to the fore.

Nonetheless, when U_P is large and monopoly profits are narrowly distributed, $\Delta C_U > \Delta C_M$. Expenditure will thus slump to

---

8. If demand for the monopolized product was completely inelastic at the expense of savings, then so long as ΔC_M was positive, total employment would increase.

the extent of the positive value of the difference $\Delta C_U - \Delta C_M$, and a stream of additional unemployment, *secondary* unemployment, will ensue. According to usual multiplier theory, instead of following this path of narrowed expenditure, knowing the aggregate marginal propensity to consume we could deduce the total fall in income and employment engendered by this contraction in total outlay. In general, income will fall until a volume of savings equivalent to the difference $\Delta C_U - \Delta C_M$ is destroyed. This will mean a fall in income and employment of those in the economic chain who had been dependent, immediately or at one or more removes, on the expenditure of the U_P group not compensated by the new expenditure ΔC_M. We have termed the sum of this unemployment secondary unemployment; to overcome it, consumption expenditures under the new income distribution would have to rise until $\Delta C_M = \Delta C_U$.

Thus we have allowed for the consequences stemming from the redistribution of income between wage earners and monopolists. Here, too, as there are differences in the propensities to consume, there are differences in savings. On the former score, of consumption expenditure, there will be immediate effects on employment; on the latter score, there will be effects on the rates of interest, unless offset by banking policy. As the degree of change in interest rates is not likely to be important for employment, we can again take the interest rate structure as fixed by banking policy. For if there is any stimulus to employment from this direction, we need hardly regard it as an aid derived from monopoly pricing; the same result could have been secured without monopoly.

DERIVED DEMAND FOR CAPITAL GOODS

Our model presupposed, implicitly at least in the discussions of rising marginal costs, an existing structure of capital equipment, and then inquired as to the effects of monopoly pricing of consumer goods output on employment. The argument can now be carried a step further, to the demand for capital goods; we are thus allowing for potential changes in the level of investment.

Ordinarily, it is held that there is a certain structure of capital goods associated with each level of employment and finished goods output. In Marshallian theory, the demand for equipment is a derived demand based on the output of final goods.

For our purposes, equipment and labor can be regarded as wanted in fixed proportions, to be in complementary demand. Therefore, insofar as unemployment is an offshoot of monopoly pricing practices, the demand for capital equipment will, through time, be narrowed. This must mean a further diminution in total income and employment, a dimming of the buoyant multiplier effects of investment-good construction. The relations need not be pursued further; they are part and parcel of the literature on the acceleration principle.[9]

MONOPOLY IN INVESTMENT GOODS OUTPUTS

So far as monopoly is present in investment or capital goods output, there seems to be little doubt that its effects are unqualifiedly baneful for employment since all installations of equipment are based on a comparison of the rate of return over cost (price) of the equipment and the rate of interest. A higher price for equipment due to monopoly will deter equipment purchases. Thus investment good activities, which are so necessary for full employment in a system with a high propensity to save, will be contracted; here too the suppressed multiplier effects will permeate the entire economy, lowering the level of employment. If the Government, through borrowing at a constant interest rate, embarked upon a public works program to restore employment lost through monopoly pricing in capital goods industries, it would have to expend a sum equal to the monopoly profits (minus any decreased rent incomes) in these industries plus the reduction in total investment goods outlay induced by the $P > MR = MC$ price policy. The one exception to the conclusion that monopoly in investment goods operates to decrease employment, an exception not deemed to be of great importance, is that of capital "deepening" — to use Hawtrey's phrase — whereby equipment is used to replace labor in output. Manifestly, monopoly pricing of equipment will make capital deepening less attractive, lowering the rate of technological displacement of laborers, thereby maintaining employment.

RECONCILIATION WITH KEYNES' THEORY

By this time the reader may have been tempted to transcribe the argument in terms of Lord Keynes' General Theory. Having

9. See e.g. P. Samuelson, "A Synthesis of the Principle of Acceleration and the Multiplier," Journal of Political Economy (December, 1939).

invaded the province of the theory of employment, it seems strange that so far we have made but slight use of Keynesian categories; on the other hand, many of our terms — idle funds, demand for securities, etc. — lend themselves to expression in MV monetary aggregates. The explanation of our neglect of the categories of the General Theory is this: Keynes' fast-flying concepts are aggregates summated for the entire system. Our structure has been more modest, for it is based upon the firm and the repercussions of its price policy. Ultimately, however, the answer as to the effect of monopoly pricing upon employment should be the same. Let us see the channels through which monopoly pricing would affect employment in Lord Keynes' Theory.

1. Taking the degree of competition as a datum, Keynes argued that the volume of employment was a unique function of effective demand: in income terms, there was a unique employment total associated with each income level.[1]

Now a change in the "degree of competition" will certainly change this function. Thus, after the shift to monopoly there is a sum of monopoly profits (and enhanced rents) at each income level; there is less employment at each income level, in amount equal to the sum of monopoly and heightened rent income divided by the average wage rate. The unemployment due to the changed employment function we termed the *primary* unemployment. It will be remembered that our income level was fixed by: (1) the constant investment volume; (2) unity elasticity aggregate consumption expenditure.

2. Keynes' aggregate propensity to consume depended upon the distribution of income.[2] Since at each aggregate income level monopolists now receive income which would have been wage income[3] prior to the institution of monopoly pricing, and which would have accrued to the primary unemployed group, we compared the propensity to consume of the one set of income recipients with the consumption propensities of the other set. Unless these are identical, there are further effects on employment. The unemployment attributable to this cause was termed *secondary* unemployment.

3. Because of monopoly pricing, prices of consumer goods are

1. The General Theory of Employment, Interest and Money, pp. 245–246.
2. Ibid., p. 245.
3. Assuming marginal costs comprised entirely of wage costs.

somewhat higher at each aggregate income level, as compared with what they would be under optimal pricing. Keynes' view was that the propensity to consume is largely a function of real income.[4] We allowed for the effects along these lines by removing, at a later stage, the initial and restrictive assumption of a fixed aggregate consumption expenditure.

4. It was Mr. Harrod who introduced the "relation" or acceleration principle into the Keynesian system. The derived demand for capital goods, according to Mr. Harrod, was dependent upon the level of output of consumer goods or, implicitly, of employment.[5] When we opened our system to include changes in the level of investment, we took cognizance of this effect: the demand for capital goods was held to depend upon the volume of employment.

5. Monopoly in capital goods industries, we argued, inevitably diminished employment. In Keynes' system this obviously fits neatly: it amounts to a reduction in the marginal efficiency of capital, and thus reduces the scale of investment.

6. Throughout the analysis, as the distribution and the level of income altered it was recognized that the liquidity preference function might shift or that, with an unchanged money supply, interest rates would be affected. This is in the Keynesian vein.[6] Also, our assumption of a banking policy designed to keep the interest rate structure from rising after the shift to monopoly pricing, or under opposite circumstances to minimize any amelioratory effects on employment of a fall in liquidity preferences, is in harmony with most Keynesian expressions. For insofar as interest rates are an obstacle to full employment, banking policy offers a ready corrective; it would be contended that there can be no real help to employment by actions which directly reduce consumption or investment. Thus effects of the shift to monopoly upon the interest rate would tend to be neglected; effects via the propensity to consume or inducement to invest would be stressed.

Hence there is nothing inconsistent between our procedure and the Keynesian system. The advantage of building up from the firm for our particular problem is that our attention at all times is focused on the relation of the firm's pricing policy to the rest of the economic system. In Keynes' analysis we must reduce the aggre-

4. Ibid., p. 91.
5. The Trade Cycle, p. 54.
6. See the most explicit statement in O. Lange, "The Rate of Interest and the Optimum Propensity to Consume," Economica (February, 1938), p. 12.

gate concepts into concrete and meaningful quantities at the firm's level. For one approaching the matter of unemployment from the theory of monopoly price our procedure should be more comprehensible. But, as shown, this is hardly to deny that all of it is compounded in the Keynesian engine.

CONCLUSION

Palpably, the view that monopoly pricing means unemployment because it diminishes output in the firm practicing a $P > MR = MC$ price policy is lacking in precision, tending either to overstate or to understate the degree of unemployment. In some segments of the economy monopoly pricing may actually increase total income and employment, a prospect not general but not inconceivable. As a final word, it is clear that our analysis pertains solely to the effects of monopoly *pricing* on employment. We took no account of possible changes in the number of firms due to monopoly pricing; moreover, there are a great many other monopoly practices which should be investigated from the standpoint of their effect on employment. Most of them would probably confirm these conclusions, although in certain cases, such as "price wars," monopoly may actually increase employment, at least temporarily.

<div align="right">SIDNEY WEINTRAUB.</div>

ST. JOHN'S UNIVERSITY

APPENDIX OF SYMBOLS

P = Price
MC = Marginal Cost
MR = Marginal Revenue
I_A = Monopoly Profits in A
ΔR_B = Increment in Rent in B
W = Average Wage Rate
ΔD_B = Increment of expenditure in B
L_C = A's employment with $P = MC$ price
L_M = A's employment with $P > MR = MC$ price
ΔL_B = Increment in employment in B
U_P = Primary unemployment
O_T = Aggregate consumer outlay
ΔO_T = Increment in aggregate consumer outlay
ΔC_U = Decrement in consumption outlay of the U_P group
ΔC_M = Increment in consumption outlay of I_A recipients
$P_1 X_1$ = A's $P = MC$ price x output
$P_2 X_2$ = A's $P > MR = MC$ price x output

GASOLINE RATIONING IN THE UNITED STATES, II

SUMMARY

The Plan: the "A" Book, 125; the "B" Book, 127; preferred mileage, 129. — Special Rations: change of residence, 130; "hardship," 132; furloughs, 133. — Fleet and Transport Rations: fleet, 133; transport, 134. — Non-highway Rations; Other Evidences: non-highway, 135; bulk and inventory coupons, 139; military acknowledgments, 141. — Definition of Gasoline, 142. — The Flowback System and the Border Problem: the flowback system, 146; the border problem, 148. — The Black Market: invalidation cycles, 149; coupon endorsement, 151; mailing centers, 153. — The Effect on Consumption, 153.

THE PLAN

Planning for a coupon system of gasoline rationing started in the late winter of 1942. There was early acceptance of the need for use of ration evidences in the form of coupons which would flow from consumer to dealer to supplier, and the predilection in favor of such a system was strengthened by experience gained through operation of the interim card plan. The card plan did not have a flowback, and a dealer could, therefore, replenish his inventory regardless of how much or to whom he had sold his gasoline. The essence of the coupon system, on the other hand, is that every transfer of gasoline, from the licensed distributor to the consumer, must be matched by an equivalent transfer of ration evidences. Further, all types of finished gasoline, with the exception of aviation fuel, were made subject to the rationing program.

The "A" Book. A minimum basic ration — the "A" book — was allowed to each registered passenger car (except fleets). It contained six series of "A" coupons, each coupon being worth four gallons, and each series having a validity period of two months. On the basis of an average of 15 miles to the gallon, the "A" book provided 240 miles of driving per month.

The basic book provided an administratively practicable method of giving sufficient mileage to each car to keep it in operating condition, and reduced the issuance burden by relieving small mileage users of the necessity for making frequent application for the rations they needed. The latter was an important administrative consideration. The local Boards had been organized for

only about six months, and they were still small and relatively inexperienced.[1] The gasoline program was more complicated and more extensive than anything they had hitherto undertaken. The issuance of a ration to each and every passenger car had also its drawbacks. Some people undoubtedly did not need their cars at all; yet it was assumed that the fact that an individual owned a car was conclusive evidence that in his opinion he needed it. Failure to issue him a ration would be tantamount to confiscation, a step which the OPA had neither the authority nor the desire to take.

An additional feature of the original "A" book was that its total monthly mileage was divided into two parts, 150 miles being deemed to be available for occupational purposes, the remaining 90 for miscellaneous family necessity driving. Behind this provision for a divided "A" book were two ideas. In the first place, a driver who was issued a supplemental ration — a "B" or "C" book — automatically had enough mileage to keep his car in operating condition and could, therefore, be squeezed on his non-occupational mileage. In the second place, it was realized that supplemental mileage would inevitably be overissued. The divided "A" book put a small mileage penalty upon the applicant for a supplemental ration, and by refusing to allow any such issuance for amounts under 150 miles per month, it operated to keep many drivers on an "A" book alone.

The divided "A" book had, unfortunately, certain defects. Persons with no occupational needs, or needs smaller than 150 miles per month, received a bonus of free driving in comparison with persons holding supplemental rations. Much more important, however, were defects which only became apparent after it became necessary to curtail the consumption of gasoline by cuts in coupon value. In November, 1942, the "A" coupon was reduced in the seventeen Eastern States (shortage area) from four to three gallons, the total monthly value of the "A" ration becoming 180 rather than 240 miles. For "A" holders with no occupational mileage, this was the end of the matter. The case of the "A" holder who also had a supplemental ration provided a more difficult problem. Was the cut in his "A" ration to be regarded as taken from the 90 miles per month of non-occupational driving, or from the 150 miles

1. They had been set up in January, 1942, to ration tires and had subsequently absorbed the automobile and sugar rationing programs.

of occupational? It seemed undesirable to take the former step and reduce the non-occupational mileage to 30 miles per month. Accordingly, the cut was in this case regarded as coming out of occupational mileage, leaving the "A" book split into 90 miles of occupational and 90 miles of non-occupational driving per month. The awkwardness of a cut, nominally in the "A" book, which affected every holder of occupational mileage, was apparent. For the first time a defect of a basic ration which mixed together occupational and non-occupational mileage was disclosed. The original "A" book, in other words, lacked flexibility. The fact that provision was made for restoring the amount of the cut, upon application, to those persons who proved need for the mileage was also unfortunate.

The second reduction in the "A" ration in the East came on March 22, 1943, when its value was halved by extending the validity period of the coupon series then becoming effective from two months to four. Once again the cut was regarded as coming out of occupational mileage. This left the "A" book, containing 90 miles per month, as a purely personal allowance, while all occupational mileage was thereafter to be issued through supplemental rations. Such a basic ration gave no mileage bonus to those holders who had no other than non-occupational driving needs. It meant, furthermore, that all occupational drivers became subject to car sharing regulations, whereas previously the "A" holder who did a small amount of occupational driving had been under pressure to share his car only because of his conscience and a desire to stretch his mileage as far as possible.

Although the exigencies of the supply situation had enabled a non-occupational "A" ration to be evolved in the shortage area on March 22, 1943, nothing of the sort happened in the rest of the nation. The delay was occasioned partly by fear of an adverse public reaction, and partly because of a failure to appreciate the technical merits of the step.

The "B" Book. The necessity for holding total driving down to the level which available supplies on the East Coast could support, and to an average of 5,000 miles per year elsewhere, dictated a reduction in the occupational driving of as many persons as possible. Occupational mileage was issued in two types of rations, "B" and "C," the difference between them being based on the amount and essentiality of mileage. To qualify for a "B" ration,

an applicant had only to show that he regularly drove to and from, or in the course of, his work, that alternate means of transportation were unavailable or inadequate, or that he shared his car.[2] The type of work which the applicant performed did not matter, so long as his needs did not exceed the maximum mileage which the "B" ration provided — originally 470 miles per month. If an applicant requested a ration larger than the "B" book, however, his occupation had to meet certain tests of essentiality, which were set forth in the Regulations.

The "B" ration provided, at the start, a maximum of 320 miles monthly, to which were added the 150 miles in the "A" book, making a total "ceiling" of 470 miles. The "B" book, consisting of 16 coupons each worth four gallons, provided a total of 960 miles and had a validity of from three months to one year. The amount of monthly mileage allowed was determined by the total life of the book. This was called "time-tailoring," and it meant that the smaller the ration the longer the period of time the ration had to run. Time-tailoring was adopted for several reasons, chief of which was administrative simplicity. In order to give the Boards the opportunity to consider carefully the applications for large amounts of mileage, the program was designed to reduce the frequency of processing of small rations. The time-tailored "B" book (with a constant number of coupons) proved, however, to be administratively awkward when coupon cuts had to be made For example, in order to issue the maximum "B" book the minimum time period had to be shortened, or else the "B" mileage ceiling had to be lowered. Both steps were taken, but both were difficult to explain to the local Boards and to ration holders.

The change away from a time-tailored "B" book was desirable for other reasons. Many people found it difficult to budget their mileage over long periods of time. Presented with a large amount of mileage at any one time, they were profligate and ran short before their books were supposed to expire. Time-tailoring became an additional handicap when supply conditions necessitated maintaining strict supervision over all rations. The long-period ration did not expire frequently enough, and therefore did not come up for

2. Each supplemental ration holder was required regularly to carry at least three persons in addition to himself, or to show that the circumstances of his employment were such (e.g. irregular hours) that such a ride-sharing arrangement could not be made.

re-examination by the local Boards. When on March 22, 1943, a second cut in the "A" book in District I had to be made, the decision was taken to drop the technique of time-tailoring and to put the "B" book upon a fixed time period.[3] A similar change was made later in the rest of the country.

Preferred Mileage. "C" or preferred mileage was always issued on a coupon-tailored basis, the ration running for a period of three months. To qualify for a "C" ration, a driver had to demonstrate his eligibility, i.e. that he performed functions regarded as essential to the prosecution of the war, and he had to show need for mileage in excess of that provided by the "B" ceiling. The latter qualification was, of course, purely administrative. If "C" coupons had been issued for all preferred mileage, no matter how small the amount, the Boards would have had to make a determination of eligibility in all such cases, and this was often a difficult matter. Moreover, very many drivers would have had to be issued both "B" and "C" coupons, when some of their driving fell in each category.

The general functions considered by the framers of the gasoline rationing Regulations as eligible for preferred mileage were five — government, education, public health and welfare, religion, and public information. But not all persons whose work is connected with the above functions are essential. The Regulations, therefore, attempted not only to limit "C" ration issuance to the most essential functions within the nation, but also to include only the most essential workers in these categories. The additional standards imposed were that the mileage must be driven directly in performance of an essential function which could be performed in no other way.

This, then, was the purpose of Section 7706 of Ration Order 5C. As originally drawn, it enumerated 20 groups or functions[4] which were eligible for preferred mileage as follows: (a) Government officers and employees, (b) School teachers or officials, (c) Transportation of groups of students or teachers, (d) Transportation of mail, (e) Delivery of newspapers, (f) Non-portable equipment for making news-reels, (g) Medical Profession, (h)

3. For technical reasons this "B" ration in District I was then issued for a period of *four* months.

4. Most of these groups had likewise been eligible under the earlier East Coast regulations (5A). A significant omission from the second list was the salesman of productive equipment.

Veterinarian, (i) Internes, medical students and public health nurses, (j) Embalmers, (k) Ministers, (l) Religious practitioners, (m) Farmers, (n) Transporting certain workers, (o) Workers in establishments essential to the community or war effort[5], (p) Representatives of organized labor and management, (q) Construction, maintenance and repair services, (r) Members of armed forces, (s) Messenger service, (t) Scrap dealers.

The actual amendments to this preferred eligibility list were surprisingly few, in view of the pressure brought to expand it. Some expansion was, of course, achieved in other ways by interpretation and by falsification of applications.

SPECIAL RATIONS

While the intent of the Regulations was that most non-occupational needs should be met within the limits of the non-occupational portion of the "A" book, it was recognized that special non-recurrent needs could not always be so covered. The "special" ration was devised to meet such needs when and to the extent that they arose. The circumstances for which special rations might be issued were clearly specified in the Regulations.[6] Most of the special rations were of little importance, and attention will here be concentrated upon those few which were of consequence either in terms of gasoline consumption or as illustrating the pressures under which gasoline rationing operated.

Change of Residence: The change of residence provision was by all odds the most important of the special rations recognized originally in the Regulations. Clearly this need for mileage was of a non-recurrent nature, and yet in the United States, with the many shifts in population occasioned by the war, provision of such

5. Three groups of essential establishments were included in this section. They were: (1) naval, military or hospital establishments; (2) common carriers and utilities; and (3) industrial, extractive or agricultural establishments essential to the war effort.

6. They included: Obtaining necessary medical or therapeutic supplies, or procuring necessary food and supplies; moving a vehicle or boat to a place of storage after repossession or upon seizure by a government authority; delivering a vehicle or boat after sale or lease; moving a vehicle or boat from one sales establishment or place of storage to another; testing a vehicle in the course of its manufacture, or moving it in the course of assembly; transporting a scientific expedition; carrying persons to and from the polls; campaigning by candidates for public office; operating a vehicle in the course of tests or experiments contributing to the war effort; demonstrating cars to prospective buyers; moving a vehicle or boat to a new residence of the owner.

mileage was necessary.[7] Accordingly, local Boards were authorized to issue special rations for a bona fide change of residence without any mileage limitation.

This liberal provision was naturally subject to grave abuse. Applicants would allege a change of residence when, in fact, they merely intended to go to a place for a limited period and then return home. A trip to Florida for the winter months was one common abuse. Transfers of members of the armed forces from post to post constituted a change of residence, and in man, cases the families of these members followed them to the new post. Workers seeking employment in new war areas applied for change of residence rations, and many of them journeyed from place to place at frequent intervals. In all of these cases the local Boards found great difficulties in distinguishing the genuine from the bogus, and no adequate solution for the difficulties was ever found. Drastic measures, such as limiting the amount of mileage to be issued or the number of bona fide changes of residences which could be recognized, seemed inadvisable. The shifts in population required by war-time demands were admittedly great, and neither the military nor the civilian agencies responsible for seeing that personnel was available where it was needed would tolerate the failure to issue the necessary gasoline. Given this situation, a considerable amount of undesirable issuance was inevitable. Efforts were made to supply the Boards with information which would help them to screen applications presented to them. Every passenger car ration holder was provided with a "mileage rationing record" on which all issuances to him were recorded. Obviously, a record which showed issuance of a large number of special rations would cause a Board to be inquisitive as to the justification.[8]

7. In Canada, however, the gasoline rationing scheme made no issuance for this purpose. Car owners who had a change of residence were forced to use their basic rations or to move their cars by public transportation.

8. Boards were instructed to ask to be shown referral cards issued by the United States Employment Service, when workers requested a change of residence ration to take them to a new place of employment. The armed forces were induced to take over issuance of rations to personnel who had been given a change of post of duty or who were travelling on official travel orders. The armed forces made their issuance in such cases through military receipts for delivery of gasoline, Form R-544. Later Form R-593 was used for this purpose. This shifted the responsibility from the local Board and also made such issuance a charge against the quota given to the armed forces, rather than against the quota given for civilian purposes.

"Hardship." Pressure from applicants, from the Congress, and from other government departments promptly supplemented the original special ration list. These supplements were both official and unofficial. The official represent those additions which were recognized by a formal amendment; the unofficial were those made by individual local Boards on their own responsibility. Many Boards, for example, felt that special rations should be authorized for visiting the sick and for attendance at funerals. In spite of the fact that no authority for such issuance could be found in the Regulations, and in spite of instructions from Washington and the Regional and District Offices which explicitly declared such issuance to be illegal, some Boards insisted upon issuance.

This conflict of opinion between Washington and the local Boards stemmed from a fundamental difference in outlook. The Boards were occasionally faced with cases of individual need which seemed irresistibly strong. An applicant might, for example, ask for gasoline to visit a sick member of his family and be able to show that public transportation was not available, that he had an inadequate amount of "A" mileage, and that on humanitarian grounds some visiting was essential. Faced by such a "hard" case, the Board would make an issuance and at the same time make a vehement appeal to Washington for an amendment which would bring its action within the scope of the Regulations. The authorities at Washington, however, were confronted with the practical impossibility of drawing up an amendment allowing special rations to visit the sick which would not open the door to gross abuses. Such an amendment, if drawn in general terms, would have led to a great increase in gasoline consumption. If, to take the other course, it were drawn strictly in an attempt to screen out the situations which did not justify issuance, the amendment would have been absurdly complicated. Faced with this dilemma, Washington for many months decided to hold the line, even though it was recognized that, as a result, some Boards would act illegally and that some, by following the letter of the Regulations, would act with severity.

Finally, however, the National Office capitulated to the extent of authorizing issuance of a special ration to cover "hardship" cases.[9] No attempt was made to define what was a hardship.

9. Amendment 96, effective February 1, 1944.

Instead, each Board was to be provided, through its District Office, with a blanket quarterly quota in terms of gallons of gasoline which it could parcel out at its discretion for issuance in hardship cases not otherwise provided for in the Regulations. It was felt that giving each Board a definite amount of gasoline for such purposes would act as a brake upon Board sympathy and generosity, even though no means of making an actual check was available.

Furloughs. One other addition to the special ration list which was the subject of much debate was that for members of the armed forces on furlough. On May 12, 1943, as a result of pressure from local Boards and the public, provision was made for issuance of a special ration to members of the armed forces on furlough.[1] The two principal limitations were that the leave had to be for three days or more and that only a flat ration of five gallons was to be used.

While these provisions were administratively simple, they were also obviously crude, and at once a demand arose that finer distinctions should be drawn. Members of the armed forces on leave from overseas service, those who had been wounded, those who had a long leave — these and many other bases were advanced as warranting a ration in excess of five gallons. A great deal of time was spent in examining, in consultation with the military, by what means the furlough ration could be improved, and finally on July 25, 1944, a change was made which seemed administratively feasible.[2] By the new amendment, one tightening provision declared that the special ration should be issued only to members of the armed forces on leave or furlough, whereas previously issuance had been permitted also to those on pass. Since passes were granted frequently, this form of provision had permitted some servicemen to secure numerous furlough rations. The new amendment further provided that the amount of the ration would vary according to the length of leave or furlough, with one gallon to be allowed for each day not to exceed a maximum of 30 gallons.

FLEET AND TRANSPORT RATIONS

The Fleet Ration. Many companies own a number of passenger cars which are used almost exclusively for business purposes.

1. Amendment 47, effective May 13, 1943.
2. Amendment 169, effective January 12, 1945.

These vehicles are frequently operated from a pool, which means that they are used interchangeably and that issuance to the individual car would destroy the pool operation. The fleet concept was designed to take care of this set of circumstances. A fleet was defined as consisting of four (later three) or more vehicles of the same type (there were also fleets of trucks, but they were made eligible for other types of rations) "owned or leased by and used by the same person or organization principally in connection with the same or related occupations." One application could be made for an entire fleet, and the ration books when issued bore not the individual vehicle registration but the fleet designation. No "A" books were issued to fleets on the ground that such cars had only occupational uses.

The fleet ration was the only exception to the basic premise of the program that each passenger car ration should be tied to the car for which it was issued. Issuance of a ration to a car (on the basis of its operator's needs) was an important control, and departure from it led to abuses. Fleet applicants listed more cars than were in actual operation, and rations issued for the idle cars were used for other vehicles in the fleet. Another problem arose because fleet operators shifted their applications from one Board to another in search of lenient treatment. These abuses were never satisfactorily remedied. At various times consideration was given to the device of freezing fleet applications at whatever Board was currently handling them. This was not done, but the Boards were instructed to question fleet operators coming to them for the first time to learn the reasons for transferring, and to secure the files from the previous Board before issuing a new ration.

The Transport Ration. The mileage requirements of commercial vehicles — trucks, busses, taxis — were originally provided for by Ration Order 5A, applicable in the seventeen Eastern States, through a special category known as the Service, or "S," ration. Since commercial consumption was not to be subject to a ceiling — that is, not arbitrarily reduced below needs — the Boards were instructed to issue as large a ration as the applicant could show that he required. The Boards were not, however, provided with standards by which they could judge the validity of an applicant's claim, and this situation, coupled with the applicant's natural desire to overstate his needs in order to provide himself with a safe cushion, resulted in a bad overissuance of "S" rations.

Partly for this reason and partly because ODT wished to be able to exercise its jurisdiction over commercial vehicles, the job of determining the size of commercial rations was passed to ODT when nation-wise rationing was ordered.

When Ration Order 5C became effective, Transport ("T") rations were issued quarterly by the local Boards on the basis of Certificates of War Necessity (CWN), which had to be obtained from ODT by all commercial vehicles. ODT by then had an organization which studied the needs of each truck, bus, and taxi, and determined the amount of gasoline needed for each quarter. This allowance was entered on the CWN, and the Board in general issued that amount without further question. Under this arrangement the local Boards had only clerical duties with respect to Transport rations.

This procedure removed from OPA the effective control of the consumption of approximately 35 per cent of the gasoline for which OPA was held accountable. After the rations for passenger cars had to be squeezed, this arrangement caused much irritation. While the OPA Boards were far from omniscient, their familiarity with local conditions would have enabled them to issue Transport rations for vehicles which had only local use more accurately than the less personal ODT District Offices. Many Boards were irritated by the necessity for issuing the certified allotments to operators without question, while at the same time they were forced to curtail passenger car mileage. ODT lacked the facilities for the continuous review which in passenger car issuance proved so invaluable. Overissuances of "T" coupons became, therefore, a serious problem. Although a number of minor technical charges were made — mostly in the direction of giving the Boards some increase of control — the difficulty was never overcome.

NON-HIGHWAY RATIONS; OTHER EVIDENCES

To cover all uses of gasoline other than for the purpose of propelling registered highway vehicles, the non-highway ration was created. "E" coupons, worth one gallon, and "R" coupons, worth five gallons, were issued for non-highway purposes.[3] At first, issuance was for a period of three months, but on March 2, 1943, as a convenience to farmers, the principal non-highway

3. They were not valid to secure gasoline for transfers into the fuel tank of a registered motor vehicle.

users, this validity period was extended to six months.[4] Non-highway rations could also be issued in bulk coupons (i.e. those for delivering to bulk storage facilities) and later in gasoline deposit certificates for ration bank accounts for the convenience of large purchasers. As with commercial vehicles, no attempt was made to limit consumption by non-highway users, in the sense that rations were cut below demonstrated needs.[5] Such consumption was regarded as essential to the civilian economy and the operation of the war program. The farmer and the construction contractor would, it was believed, only buy gasoline in such amount as was necessary to do their jobs. Furthermore, the nation-wide gasoline rationing program rested upon conservation of tires and non-highway consumption of gasoline did not, with minor exceptions, use tires. Issuance of coupons was required only to prevent the leaks and diversions to curtailed uses that would otherwise occur. The local Boards, therefore, in processing applications for non-highway purposes tended to issue the amount that the farmer or contractor said he needed. The result was a great overissuance of non-highway coupons. Applicants estimated their needs with liberality, usually with no idea beyond being on the safe side. If the surplus issuance had been turned in to the local Board each time a new ration was issued to a farmer (as the Regulations required), or even if it had been allowed to remain unused in a bureau drawer, no great harm would have been done. Unfortunately, the overissuance tended to be given away or sold, particularly to gasoline dealers. How could issuance be tightened without having a detrimental effect on food production or raising a storm of protests from the farmers? How could overissuance be neutralized, if it occurred?

The first attempts to secure a more exact issuance were tentative and fumbling. Local Boards were advised to add farmers, contractors, etc. to their panels which passed upon gasoline issuance in the hope that thereby more expert knowledge of needs might be secured. Another step was the creation of County Farm Transportation Committees which were to certify applications of farmers for non-highway rations to the Boards. OPA also under-

4. It remained three months for non-agricultural users.
5. In the cases of non-occupational use of boats, and in the case of amusement devices, ceilings were set which were relatively low. Quantitatively these uses of gasoline were unimportant.

took a thorough revision of the non-highway application form and of the instructions to local Boards concerning non-highway issuances. The rate of consumption of the principal types of non-highway equipment depends upon so many factors that no one simple guide is adequate. However, tables were prepared showing probable gallonage consumption per hour, under specified conditions, of all important types of tractors, as well as figures showing average annual use of tractors in various areas.

Three other more elaborate schemes were tried out. The first seized upon the proposition that most non-highway gasoline was delivered to the farmer in bulk to his premises by tankwagon. It seemed reasonable, therefore, to prohibit transfer of gasoline in exchange for "R" (five-gallon) coupons except in bulk. This meant that retail outlets in cities which did not make such transfers would have no right to have "R" coupons in their possession, and that therefore many of the illegal markets for "R" coupons would be sealed off. "E" (one-gallon) coupons were, however, to be acceptable at all retail outlets in order to take care of the few farmers and others who did not get deliveries in bulk. In addition, holders of "R" coupons who might be in a similar position were entitled to go to their local Boards and exchange "R" coupons for "E" coupons upon showing need.

Although this scheme was adopted (April 1, 1944) after the most careful deliberation, trouble arose at once in Kansas, Oklahoma, and North Dakota. In these States the practice of delivering in bulk to the farmers was much less common than had been estimated. Instead, many farmers were accustomed to come to retail outlets in order to take delivery in drums. Dealers who did this kind of business insisted that the amendment would drive their customers away from them, since these customers, instead of exchanging their "R" coupons for "E" coupons at the local Boards, would prefer to go to the dealer who, because he made bulk deliveries by truck, was entitled to accept "R" coupons. So vigorous and so genuine was the complaint that amendments were passed which, in effect, took the three States mentioned above out from under the operation of the plan. A substitute was devised, but it was so loose as to be ineffective. Five months later the scheme was completely abandoned.

A second scheme, tried out in a portion of Minnesota, eliminated issuance of non-highway coupons to farmers who had bulk

facilities. Instead, purchase receipts, similar to checks, were provided. These evidences, after passing through the flowback, were sent back to the Boards in order to be checked against the issuance made to individual farmers. The scheme proved, however, much too elaborate for effective use and it also was abandoned.

A third scheme, tried out in Lancaster, Pennsylvania, stemmed from the simple premise that issuance of non-highway rations could be checked against and based upon actual records of past consumption. Every applicant for a non-highway ration was provided with a delivery record (Form R-585) upon which was to be entered the gallonage value of the *issuance* made to him. Subsequently, *all bulk deliveries* were to be entered on the record by the supplier. When the farmer applied for a new ration, his record had to be sent to the Board. This record of actual past consumption would obviously provide the Board with a basis upon which to base a new issuance, and it would, moreover, indicate to the Board how accurately past issuance had been made. After the Board had determined the gallonage the farmer required as his new ration, it was to issue him his coupons in this amount, *minus* any remainder of the coupons issued to him previously and not accounted for as used on his record. In short, this remainder was to be left outstanding in his possession and regarded as available for the period. Such a provision was designed to mop up past over-issuance, while the record itself was designed to secure more accurate issuance for the future. The plan was put into effect in the limited area of Lancaster County on March 10, 1944, and favorable reports concerning its operations were received. It was made nation-wide on May 1, 1945.

The outcome of all of these approaches, together with a greater awareness by Boards and consumers of the need for conserving gasoline, was a gradual improvement in non-highway issuance. But the result was never satisfactory and, although statistics cannot be given, some of the excess of non-highway coupons flowed to the black market. No arbitrary cut in the value of the non-highway ration was ever made, principally because of fear of political repercussions. In addition, an equitable method of making a cut could not be worked out. Reduction in the value of "E" and "R" coupons alone would have left untouched the rations, or portions of rations, issued in the form of bulk coupons or as bank accounts.

Even with the benefit of experience, it is not easy to see how a satisfactory scheme for rationing non-highway gasoline could be devised. Without extremely cumbersome controls, over-issuance of such rations is inevitable. The farmer cannot estimate his needs with precision because what he will use depends upon the weather — an incalculable variable. The farmer cannot be expected to take a minimum necessary amount of coupons with the assurance that, in the event of a deficiency, he can apply for and be given more, because he has not as quick access to a local Board as has the town dweller. Any attempt to tighten the issuance by prescribing tighter control through application forms which ask for detailed information runs up against both the inability of many farmers to supply the information and their resistance to bureaucratic administration.

Bulk and Inventory Coupons. In establishing coupon types and values, a conscious effort was made to keep the majority in small denominations. The possibilities of racketeering indicated the advisability of making that occupation as unprofitable as possible, and large denomination coupons would provide an easy medium for use in the black market. Two large denomination coupons seemed, however, to be necessary. (1) Inventory coupons, issued to dealers to cover the difference between their storage capacity and their actual supply of gasoline on hand at the time of registration, were of 100-gallon denominations (with one-gallon inventory coupons as "change"). (2) Bulk coupons were also in 100-gallon and one-gallon denominations, because purchases in bulk were bound to be of large gallonage. To require inventory and bulk transactions to be covered by five-gallon coupons would have been an exhausting waste of time and paper.

The more important of these 100-gallon coupons — and the more troublesome — was the bulk coupon issued to large consumers of gasoline. Any person who could prove that he had bulk storage facilities from which he serviced his vehicles or equipment was entitled to have his ration issued all or in part in bulk coupons, or to have his ration book marked for bulk transfer.[6] In order to spare the Boards some inconvenience where that would work no hardship on the consumer, a minimum monthly ration of 250

6. In general, all highway coupons were good only for transfers into the fuel tank of the motor vehicle specified on the cover of the ration.

gallons — later raised to 480 gallons — was also established as a prerequisite to bulk coupon issuance.

Since any large consumer was eligible for bulk coupons, this meant that some "B" rations, some "C" rations, some "T" rations and some "R" rations were issued in bulk coupons. The original identity of these rations was lost, and no way existed by which to determine the particular type of consumption for which the ration had originally been issued. Moreover, a cut in the value of the bulk coupon could not be made for any one class of consumers without reducing the rations of all other holders. Similarly, a cut in the value of passenger car coupons, or "T" coupons, or "E" and "R" coupons, left holders of bulk coupons untouched. Since only a very few fleets had bulk coupons in lieu of passenger car ration books, this was not serious as regards "B" and "C" coupons. However, 20 per cent of non-highway and nearly 40 per cent of total transport issuance was made in bulk coupons. Because both transport and non-highway rations were badly overissued, bulk coupons were also badly overissued. In the summer of 1943 a system of ration banking was established nation-wide. Licensed distributors and large consumers were given gasoline accounts with designated banks, and bulk coupons were abandoned.[7]

Inventory coupons, like bulk coupons, were issued in 100 and one-gallon denominations. When, at the beginning of rationing, all dealers registered at local Boards, they were required to state their total storage capacity and the amount of gasoline which they currently held. In order to permit them to fill their tanks on subsequent delivery, inventory coupons of a gallonage equal to the amount of unfilled storage capacity were issued to them. A special coupon was needed for this purpose, since purchase of gasoline for inventory did not constitute consumption. Only when this gasoline passed to consumers, and coupons were collected from them in exchange for it, was it "consumed" within the meaning of the gasoline rationing program.[8] The size of inventory coupons made them a lucrative counterfeiting proposition, so much so that

7. Amendment 50, effective May 15, 1943. As a convenience to farmers the 100-gallon bulk coupons were allowed to remain valid until August 31, 1943. The one gallon bulk coupons continued valid until December 31.

8. Issuance of inventory coupons was also authorized to cover dealer losses resulting from evaporation, spillage and theft of gasoline and coupons. Dealers also received inventory coupons in place of invalidated evidences which had been legally collected but could no longer be used.

they eventually had to be cut off. This was not done, however, until the summer of 1944, when local Boards were authorized to open ration bank accounts and the certified ration check could be substituted for the inventory coupons.

Military Acknowledgments. While the military was, of course, not subject to civilian restrictions, it was necessary to provide some evidence for military use when gasoline was purchased from an outlet otherwise held responsible for rationing regulations. So long as military withdrawals occurred above the licensed distributor level, they were not accountable. Purchases below that level, however, were accountable by ration evidences.

The problem was to provide the dealer or other seller with some evidence that he had in fact sold to the military authorities. This was accomplished by use of an Acknowledgment of Delivery (Form R-544) which resembled a check and which was presented for replenishment just as were coupons. The military Acknowledgment was intended primarily to provide for emergency needs, for example, the •cross-country movement of military vehicles. It soon came to be used, however, by military personnel travelling in private cars. Had issuance been carefully handled, this would have caused no difficulty. Actually, issuance was extremely loose. Not only were large numbers of forms issued, but they were often issued in blank. They could be cashed for any amount, and often the service station operator filled in the amount himself.

The form was changed twice in an attempt to control its use. Finally, on October 1, 1944, it was invalidated altogether, and a new and stricter system provided for military use. Two forms were issued for transfers into the fuel tanks of vehicles, one (R-593) for vehicles owned by the armed forces, the other (R-594) for private cars driven by military personnel in pursuance of their duties. The former had no specific value, the latter was good for only five gallons per form. Military establishments buying gasoline in bulk had to open ration bank accounts. In addition, the military authorities themselves took steps to assure more careful issuance of these forms.[9]

9. Two other unimportant forms were the *exchange certificates* and the *emergency transfer receipts*. The *exchange certificate* was used, before the introduction of ration banking, to provide the licensed distributor with a summary evidence for use in making his monthly reports to the State Tax Administrator and to OPA. To send all the gummed coupon sheets he had accumulated during the month would have been awkward and futile, since inspection of

DEFINITION OF GASOLINE

The gasoline rationing program was, from the start, designed principally to control the use of gasoline in motor vehicles and particularly in passenger cars. In defining the product to be rationed, therefore, the chief test was use to propel a motor vehicle by means of an internal combustion engine. There is, however, a wide range of petroleum products which, alone or in blend, will propel a motor vehicle. Many of these products — and they were quite different in different areas — which had not been commonly or usually used before rationing, suddenly came into common and usual use. Should coupons be required for transfers of these fuels? Different dealers and distributors answered this question in different ways.

The fuels which caused the most trouble under this clause were the so-called "tractor fuels." These tractor fuels were expressly designed for sale to farmers in states permitting no tax refund or exemption on gasoline used for non-highway purposes. Their specifications were carefully adjusted to escape the range of the state definition of gasoline for tax purposes. Since a tractor was not a motor vehicle according to the definition of the latter in Ration Order 5C, the tractor fuels were not gasoline. On the other hand, demand for these tractor fuels was stimulated after rationing, in those States in which they were available, because some of them actually could be used to propel motor vehicles. When coupons were collected for transfers of these fuels, coördination with the State tax reporting systems — which excluded the fuels — was upset. Even their legitimate use in tractors, without coupons, caused difficulty, since farmers could apply for and receive rations

them would have been impossible. The coupons were surrendered, instead, to the local Boards and exchange certificates which amounted to receipts were issued in their place.

Emergency transfer receipts were provided for emergency use when inability to secure gasoline promptly would have constituted a serious threat to life, health or valuable property. The forms, which had no specified value, were distributed to service stations. A person needing emergency gasoline had to fill out his name, address, and the purpose for which the gasoline was needed. The dealer then took the form, which was in duplicate, to his local Board, receiving inventory coupons in return. The Board filed one copy for its own records and sent the other to the transferee's Board.

The emergency receipt was a source of much petty fraud and most of the forms were withdrawn from the filling stations by the District and Regional Offices They were not invalidated, however, because of legal insistence that equity would be damaged thereby.

to run their equipment which was in fact operated on non-rationed fuels, thereby leaving a surplus issuance for diversion to highway uses.

This situation was further complicated when, with the advent of fuel oil rationing, it became necessary to draw some dividing line between gasoline and fuel oil for rationing purposes. The Gasoline and Fuel Oil Branches of OPA early agreed that between them almost the whole range of petroleum products should be covered, since otherwise the unrationed fractions could be readily obtained and blended up or down to make motor fuel or fuel oil, thereby defeating the purposes of both programs. As to the line between them, a 100° flash point seemed the best answer, and the gasoline definition was amended accordingly to include all fuels (with certain exceptions to be noted) having a flash point below 100° Fahrenheit. It was admitted that this would throw into the gasoline range some products that were really not like gasoline and which, therefore, if excluded, would not endanger the gasoline program. On the other hand, these products were highly explosive. The Fuel Oil Branch was understandably reluctant to permit sale with fuel oil evidences of any product which, if used as a burning oil, might endanger life and limb. As a result, certain products were included as gasoline which were later to cause trouble.[1]

The new definition of gasoline, therefore, became any petroleum product "commonly known or sold as gasoline" or having a flash point below 100° Fahrenheit, with certain exceptions of which the most important was naphtha.

Beginning with Ration Order 5C, naphtha was always excluded from the gasoline definition. It was readily admitted from the beginning that such exclusion presented opportunities for evasion. On the other hand, the alternative of rationing naphthas presented numerous administrative difficulties which might be more costly than the potential leak. For example, many naphthas are sold for a variety of home purposes (cooking, lighting, cleaning)

1. The same amendment which introduced the 100° flash point made other important changes. The phrase, "commonly and usually used to propel a motor vehicle," was changed to read "commonly known or sold as gasoline." Further, all petroleum products, excepting liquified petroleum gases, including all those products otherwise excepted, became gasoline when they were used or blended for use in a motor vehicle. Products like the much-debated tractor fuels were in this fashion clearly included by the language, and at the same time the door was at least partially closed on the use of unrationed products in blend.

in small quantities by hardware, drug and grocery stores. Inclusion of these products in rationing would, it was estimated, have necessitated the registration as dealers of approximately 300,000 such outlets — a formidable local Board work-load. There was also reluctance to include these minor uses, because of the volume of coupon issuance it would entail. Unless naphthas were rationed on separate evidences — a big printing and distribution job in itself — gasoline coupons, principally non-highway, would have to be used, and coupons issued for purchase of naphtha would, therefore, be equally good for purchase of gasoline. In addition, establishment of general eligibility for a ration would mean that a great many people would apply for it. Therefore, actual encouragement of naphtha consumption (and for illegal purposes) might result.

This attitude underestimated the potentialities of naphtha as a substitute motor fuel. Trouble soon arose, particularly in the Middle West. The Regulations had made no attempt to define naphtha, because this would have been a highly complicated definition. For practical purposes, therefore, dependence was placed upon trade practice to provide a definition. Unfortunately, the same product was sold as gasoline by one company and as naphtha by another. Thus the anomalous situation existed that a product was rationed in the first instance and excluded in the second. And all that could prevent abuse was the good faith of the industry and the use clause which made a rationed product one which was used or blended for use in a motor vehicle.

When it became apparent that OPA's trust in the good faith of some sections of the industry was misplaced, an attempt was made to reinforce the position taken under the Regulations. To this end, it was forbidden to deliver naphtha as an unrationed product into the fuel tank of a motor vehicle; to store naphtha in tanks registered as a part of gasoline storage capacity; to receive coupons in exchange for sale of products bought without coupons as unrationed naphtha. It was also provided that dealers and distributors might not sell unrationed naphtha to a customer who they knew would use it as a motor fuel.

These steps were ineffective. The Standard Oil Company of Indiana, in particular, had for a long time been marketing as stove and lamp naphtha a product which was really unleaded third-grade gasoline. Definition of naphtha by trade name excluded this product from rationing. Standard of Indiana was not blind to its

opportunities, and pushed the sale of this product. Soon other marketers were forced to come out with competing products. Abuse became so widespread that further action was clearly indicated, preferably without resort to rationing. In October, 1943, PAW issued a limitation order[2] which, with certain exceptions, limited deliveries of naphtha to dealers and bulk consumers during any calendar quarter to 125 per cent of the amount they had received during the comparable quarter of 1941. Industry figures indicate that the limitation order had some effect during the three months following its adoption. However, in March the sales trend moved upward again.

On June 29, PAW asked OPA to ration naphtha in certain Midwestern States.[3] Following a series of conferences in which the problem was explored, and after industry and OPA field opinion was found to be overwhelmingly in favor of the move, OPA agreed to accept an appropriate directive from WPB. The directive was issued and an order was prepared rationing naphtha in all of PAW District II.[4] The amendment[5] became effective November 20, 1944.

Hopes that the naphtha problem would be solved by this move were quickly dashed. As indicated above, the prospect of rationing naphtha was not entirely attractive to OPA, because of its implications in terms of Board work-load and excessive coupon issuance. In an attempt to exclude as many specialized products as possible, the definition was purposely made narrow. It soon developed that it was too narrow. Not long after the order was issued, the Standard Oil Company of Indiana brought onto the market a naphtha with a 10 per cent point of 230° F. This was an entirely acceptable fuel for the limited automotive use for which it was intended, and the abuse continued.

2. Petroleum Distribution Order 18.
3. Illinois, Indiana, Michigan, Minnesota and Wisconsin and "in any other states which may be agreed upon by OPA."
4. North Dakota, South Dakota, Nebraska, Kansas, Oklahoma, Minnesota, Iowa, Missouri, Wisconsin, Illinois, Michigan, Ohio, Indiana, Kentucky, and Tennessee.
5. Amendment 159. The move was accomplished, not by general amendment of the gasoline definition, but by the addition of a clause including naphthas "under circumstances set forth in this order." (Paragraphs 775–780, Loose Leaf Service.) These fuels were then separately defined as "having an ASTM 10 per cent distillation point below 220° Fahrenheit." Non-highway coupons were issued for their purchase.

THE FLOWBACK SYSTEM AND THE BORDER PROBLEM

The Flowback System. Rationing is a process of controlling demand for a scarce commodity. Therefore; the first step in the functioning of a rationing program must be the placing of ration evidences (representing rights to buy) in the hands of the right consumers and in the right quantities. But there is more to the problem than issuance of ration currency. In order to control effectively the use of that currency, and the distribution of the commodity it represents, supervision must be exercised over the industry engaged in distributing the commodity. In the gasoline rationing program, the principal control factor was the flowback system.

In distributing gasoline, the organization of the petroleum industry falls roughly into three levels. At the top stands the licensed, or primary, distributor, licensed by the State or States within which he operates to import gasoline and to account for taxes. Next stands the intermediate distributor, who acts as a wholesaler. His distinguishing characteristic is the fact that he must buy his supplies from a licensed distributor. At the bottom of the distribution system is the retailer, who may operate a service or a tank truck, making deliveries to consumers only.

Gasoline flows from the licensed distributor downward. Coupons, on the other hand, entered the distribution system at consumer level and flowed upstream in gallonage amounts equivalent to the downstream movement of gasoline. Throughout the system, each transfer of gasoline had to be matched by a like transfer of coupons. Thus there was, at every level, a control over the movement of rationed product.

Specifically, the flowback of ration evidences worked in the following manner.[6] The consumer received his coupons from a local Board. If he made large purchases of gasoline in bulk, he received a gasoline deposit certificate[7] which he deposited in his ration bank account. When he bought gasoline from his dealer, he surrendered coupons (or a ration check) of a gallonage value equivalent to his purchase. The dealer was permitted replenish-

---

6. Ration banking will be included in this description It was installed nation-wide on January 27, 1943. While the system as first set up did not include it, the theory of flowback was not thereby changed. The original provisions will be briefly described in subsequent footnotes.

7. Bulk consumers at first received bulk coupons.

ment only to the extent of his consumer coupon collections.[8] When he received a delivery from his supplier he surrendered the appropriate number of coupons. For convenience in handling, these were placed upon gummed sheets,[9] each holding fifty coupons of a like type and value.

The supplier who accepted these gummed sheets deposited them in his ration bank account.[1] If he were an intermediate distributor, he had to write a ration check to his supplier to cover deliveries made to him. If he were a licensed distributor, he had to write a monthly check to OPA for the amount of his rationed sales during the month. With this check the licensed distributor had to file a reconciliation statement (Form R-550), accounting for all sales of gasoline made during the month. He had also to prepare a report form, for tax purposes, for submission to the Tax Administrator in each state in which he sold gasoline. To the latter were attached the OPA report and check. These received preliminary auditing by the Tax Administrator. The R-550 and ration check were then forwarded to the audit and control section of the Gasoline Rationing Branch, Washington, where they were finally checked.

The advantages of the tie-in of ration reports with State tax reporting procedures were numerous. The states depend heavily upon the revenue received from motor fuel taxes. To assure their collection, extensive controls over licensed distributor operations are maintained. The leakage which might otherwise be possible between the refiner and the distributor was largely plugged by strict state auditing systems. At the refineries, and at all points of entry (e.g. pipelines and water terminals), the states require the use of approved entering devices, sealed marketing tanks, and separate loading facilities for shipments by various types of transport. Subsequently, the States secure reports of carrier operations of all types from the carriers themselves and from the

8. Except for inventory coupons issued to him upon registration to the extent of his unfilled storage capacity.

9. Form R-120.

1. Prior to ration banking, the gummed sheets were surrendered to the Boards who issued Exchange Certificates (R-560) in return. This was a device to reduce the volume of paper handled at the upper levels of operations. It also brought the coupons back to OPA at a level where they might be checked. Actually the volume of paper, even at the Board level, was too great to permit effective scrutiny of the evidences.

Department of Interior. Exchange of information among States enables each to follow the flow of gasoline from the refinery through every level of distribution to the point of tax payment.

In addition, the States maintain sizable forces of inspectors and auditors to check all sales made within their borders. These audits are made with relative frequency. The Federal tax system, on the other hand, which is imposed at the refinery level, is usually three to four years in arrears. This was an important factor governing the decision to tie-in ration reporting with the State tax systems, rather than to carry it up to the refinery level, where the coverage would have been considerably wider.

A drawback of the State system was the large number of reports that had to be handled. State licensing systems differ widely, and as a result in some States relatively few distributors are licensed, while in others, such as Kansas, hundreds of small operators are required to file reports. On the other hand, the coöperation received from tax administrators in making preliminary checks and the convenience to the industry of coördinated reporting could not be overlooked.   •

In addition to the controls provided by the flowback system, OPA maintained some supervision over the industry through registration. Before rationing began, each dealer and distributor was required to register with his local Board. Dealers and intermediate distributors were required to report their total storage capacity and their total gasoline inventory as of the start of rationing. Inventory coupons were then issued to the amount of the registrant's unfilled storage capacity. At no time could he have on hand gasoline or coupons in excess of his registered storage capacity. The debiting program, which began early in 1944, made good use of this clause by making penalty reductions in registered storage capacity equivalent to the gallonage value of invalid evidences surrendered.

The Border Problem. The effectiveness of gasoline rationing was impaired in irritating, although minor, ways by what can be called the border problem.[2] The existence of an unrationed area next to a rationed one, as was the case when the East was rationed

2. During the period before the program was nation-wide, a "no sale" and a "buffer" zone were created west of the rationed area, and these gave rise to minor inequities which were literally insoluble. Even more troublesome were problems arising from deliveries from the rationed to the unrationed area by pipeline, tank car and truck.

and the rest of the country was not, obviously caused difficulties, and these were perpetuated by the emergence of the Eastern shortage area. The fact that a coupon was worth less in terms of gallons on one side of a border led drivers to spend on the other side of the border. Such a situation, while creating a zone within the shortage area in which the force of the rationing system was ineffective, did not seem worth attempting to correct. More important was the impact of the difference in coupon values upon dealers and their supplies. A dealer who only gave three gallons per coupon which his competitor across the border accepted for four gallons was injured. Suppose, however, that this dealer received deliveries from a supplier outside the shortage area. Under such circumstances, a number of avenues of fraud were opened to both. The supplier could sell only three gallons per coupon to the dealer, redeem the coupons at his bank for four gallons, and distribute the bonus as he saw fit. Or, entering into collusion with the dealer, he could sell to him at four gallons per coupon, after which the dealer had a bonus which he could distribute on the black market.

The enforcement technique finally evolved after much debate was to require dealers to mark clearly on the gummed sheets, to which coupons were attached before being passed on to suppliers, the name of the county in which they were located. These sheets, when deposited by the supplier in his bank, were to be inspected by the bank to see where sale of gasoline was made, with credit given at the appropriate value per coupon.

The Black Market

The black market did not become prominent until the gasoline rationing program was six to nine months old. For one thing, except in the East, the program did not bear heavily on users of gasoline and therefore no strong incentive existed for growth of a black market. For another, both the consumers and the industry were unacquainted with how the ration machinery operated. They did not know how, even if they had the desire, to engage in illegal operations with a minimum of risk. Both of these safeguards were lost with the passing of time. When it became necessary to cut passenger car rations, a demand for black market gasoline was stimulated, and by this time very many people knew how the Regulations could be evaded with relative impunity.

Those who were responsible for framing and operating the program had stressed the point that the best safeguard is to secure correct issuance. It is a truism to say that, if each applicant secured only enough coupons to cover his legitimate needs, no surplus would exist which could be diverted to illegal use (although counterfeiting might grow up). The difficulty of securing proper issuance was not appreciated until early records came in which enabled a comparison to be made of issuance and flowback. In the case of non-highway coupons and "S" coupons (the predecessors of "T"), the figures were shocking. A great excess of unspent coupons existed which hung over the market and which represented potential consumption. It was clear not only that a drive had to be made to improve issuance, but also that steps had to be taken for periodical alteration of the type of coupons, so that the over-issuance of a prior period would be invalidated. The tightening came in various ways. Starting in the summer of 1943 it became the practice to have a regular cycle of invalidation of old and replacement with new coupons. As local Boards, through experience, became acquainted with the purpose of the shift in the type of coupons, as ration banking was extended, and as more vigorous enforcement took place, it was possible to erase the overhang of unused coupons which had been issued in an earlier period.[3] An additional advantage of periodic invalidation was that it withdrew a particular type of coupons from circulation before counterfeiting techniques could be perfected; and by making the validity period of the coupon relatively short, it made counterfeiting less profitable.

In December, 1943, the "T" coupons for the first quarter of 1944 were issued in strips rather than in books, and were serially numbered. Three months later the same technique was used for "B," "C," "E," and "R" coupons. This was an important device from the enforcement viewpoint because the serial numbers rendered the job of the counterfeiter much more difficult. It was important also because the serial numbers provided identification of the coupons with the ration holder, which made them easier to trace and easier to identify by the station operator. Another

3. Some local Boards, not at first appreciating the purpose of the cut-off, continued issuance of old coupons beyond the time specified, with the result that some ration-holders found themselves innocently in possession of invalid coupons.

important advantage of the strip technique was the elimination of tailoring. Coupons which had been torn out of books in the tailoring process were technically invalid. But they often found their way into black market channels, for theft was relatively easy and identification and accountability difficult.

Another technique designed to check the transfer of surplus coupons was coupon endorsement. This was introduced on November 16, 1942, while gasoline rationing operated only in the East. It was carried over into the National program and for many months was the subject of bitter attacks by the industry and the public. Partly as a result of its unpopularity and partly because the full potentiality of coupon endorsement as a weapon for enforcement was not realized, this requirement of the Regulations was long ineffective. Endorsement was also often executed by the dealer rather than by the consumer of gasoline, and spokesmen for the dealers insisted that to expect dealers to demand that their customers endorse coupons in advance was to fly in the face of established business customs. At various times the protest was so intense that the requirement of endorsement was on the verge of being knocked out of the Regulations.

The phenomenon which put life into the technique of endorsement was the emergence of a large black market — and especially a black market in counterfeit coupons. It should be appreciated that the black market from the outset operated almost entirely within the coupon system.[4] The most common method was for a dealer to acquire, by purchase or by gift or in some other way, an excess of coupons. With this excess he could sell gasoline — at a black market price — to customers who did not turn over coupons of their own to cover the sale. The coupons which the dealer had in possession did, however, get into the flowback, and these therefore used up some of the gasoline which had been allotted by PAW for legitimate civilian consumption. So far as the illegal consumption grew, the portion left for legal consumption was diminished.

The counterfeit situation arose as criminal elements came to see that rationing provided them with the chance for a profitable

4. One exception was the so-called drip gasoline, which was taken illegally from pipelines. This was an old practice, existing long before rationing. Obviously such gasoline never got into the machinery of the gasoline rationing system.

racket. The early counterfeits were relatively crude, both in their make-up and in the type of paper on which they were printed. Soon, however, the racketeers managed to secure safety paper and to produce counterfeits which could only be detected by fluorescent lamps. This was counteracted by limiting and safeguarding the plants which did the printing job, and by inspecting their procedures to prevent theft.

The job of searching out counterfeit coupons which were in the flowback proved beyond the resources of the Boards. It was then shifted to the ration banks and to the District Offices. Finally, in the spring of 1944, a verification center was set up in each of the eight OPA Regions, equipped with fluorescent lamps and other scientific devices for detecting counterfeits, and staffed with trained personnel. All coupons were sent from the banks to them, and in this fashion a 100 per cent coverage of all evidences entering the flowback stream was achieved.

In the great bulk of cases counterfeit coupons were sold to filling stations rather than to the public. Such a sale was obviously less risky, since it would be for a considerable gallonage. Moreover, the public appreciated that purchase of coupons was illegal, as it often did not appreciate the illegality of purchases of gasoline without coupons. Because of this fact, enforcement focused attention on the filling stations. Obviously dealers were unlikely to get many counterfeits if they insisted that a motorist, when buying gasoline, transferred properly endorsed coupons from a ration book. Motorists who were in possession of illegally secured coupons would naturally be timid about opening themselves to detection by placing endorsements on such coupons. If the dealer innocently did take counterfeit or stolen coupons, he was charged with them precisely as though he had done so deliberately, i.e. he was notified that his registered storage capacity (his operating inventory) would be reduced by the gallonage value of the bad evidences. It was felt that in most cases insistence upon endorsement and a check of the serial numbers of the coupons presented against those appearing on the consumer's ration folder[5] would provide sufficient safeguards against acceptance of bad coupons from consumers. Obviously this system was bound to do injustice to some dealers as

5. The issuing Board noted on the folder the inclusive numbers of the coupons in each ration.

counterfeiting techniques improved, and in December, 1944, a limited relief was authorized.[6]

In addition to counterfeiting, black market operators stole legally printed coupons from Boards, warehouses, etc. Thefts were difficult to prevent, because in the earliest days of the program, Boards had inadequate facilities for safeguarding coupons. As a result Boards were restricted in the size of the coupon inventory which they were permitted to hold, and they were instructed to keep the bulk of their inventory in a security site, such as a bank.

Most important of the safety techniques, however, was the mailing center. Boards served by a mailing center held a minimum of coupons. Instead, coupons were held at a central location, where they could be given adequate protection. The Boards notified the mailing center of the proper issuance to make, and the center prepared the rations, wrote the required information on the identification folder, and mailed the ration to the individual. The Boards themselves could issue, in case of emergency, only purchase permits. These were relatively safe from theft because they were difficult to forge and so were unprofitable for black market purposes.[7]

THE EFFECT ON CONSUMPTION

The ultimate test of the accomplishment of gasoline rationing is its effect in reducing consumption. As evidence of this reduction the estimates of motor-fuel consumption, 1941-44, prepared by the Public Roads Administration, are given in the table below.

These figures indicate that civilian gasoline consumption in 1943 was 71 per cent, and in 1944 was 73 per cent, of 1941. The most significant figures are, however, those for passenger cars, which show consumption in 1943 as 60 per cent, and in 1944 as 62 per cent,

6. Amendment 163.

7. Another difficult problem concerned the transfer of a gasoline business to a new operator by a person who had been suspended or debited out of business, or who was short of evidences as a result of selling gasoline without collecting coupons. The drastic solution would have been to apply all penalties to the business location itself, as well as to the operator. If the station was suspended, it would have to stay closed for the duration of the suspension, regardless of who operated it This would, however, penalize the innocent owner of the property by completely depriving him of income from his investment during the suspension period, and the padlocking procedure was not adopted. As a result, loopholes remained, especially with respect to leased stations, which successive tightening amendments did not remedy.

CIVILIAN MOTOR-FUEL CONSUMPTION[8]

| | 1941 | 1942 | 1943 | 1944 | 1941 | 1942 | 1943 | 1944 |
	(Thousands of barrels per day)				(Percentages)			
Passenger cars...	1,172	940	706	734	100	82	60	62
Commercial.....	406	358	338	340	100	88	83	84
Non-highway....	146	163	172	178	100	112	118	122
Total........	1,724	1,461	1,216	1,252	100	85	71	73

of 1941. As has been explained, gasoline rationing did not attempt to reduce consumption by commercial vehicles and for non-highway purposes. The farmer was provided with all the gasoline evidences for which he could show occupational need, and non-highway consumption actually increased by 22 per cent in 1944 over 1941. Commercial use of gasoline decreased somewhat, partly because the number of trucks and busses shrank by 7 per cent and partly because of rationing restrictions.

Not all of the decrease in passenger car consumption is, of course, attributable to gasoline rationing. The number of such cars in operation dropped by 17 per cent in 1944 over 1941, and indirectly tire rationing had some effect in reducing gasoline consumption. Against the weight of such factors one should, however,

8. These figures, which in detail have not been published, were prepared for me by Mr. G. P. St. Clair. The breakdown does not coincide precisely with that used in gasoline rationing. For example, the figures for passenger cars given above include consumption by taxicabs, whereas in the rationing plan this consumption is classified as commercial.

It may be asked why the effect of gasoline rationing is not indicated by figures of gasoline flowback. The difficulty is to tie in such figures to the figures for a pre-rationing year. This difficulty arises not so much from the overall figures as from the distribution of these figures among the chief types of use — passenger cars, commercial, and non-highway. The flowback figures in 1943, for example, total approximately the same as those given by the Public Roads Administration. But the flowback figures for commercial and non-highway uses are much less than those shown by the Public Roads Administration. What is the explanation? In part it arises because gasoline was secured through non-highway and commercial coupons, and used in passenger cars. This black market consumption is reflected in the non-highway and commercial flowback, and it may amount to 100,000 b/d. It is, however, probable that the Public Roads Administration figures underestimate non-highway and commercial uses in wartime. Any attempt to reconcile flowback and Public Roads Administration figures would be a complicated statistical job. For present purposes, therefore, the Public Roads Administration figures are accepted with the general warning that they may underestimate the reduction in passenger car consumption which gasoline rationing accomplished.

consider the effect of expanding incomes during the war period. Although this factor cannot be measured, its significance in increasing the demand for gasoline cannot be doubted.

Another statistical measure of the effect of gasoline rationing is provided by the estimates of motor vehicle travel prepared by the Public Roads Administration. They show that vehicle-miles travelled by passengers in 1943 were 58.5 per cent, and in 1944 60.8 per cent of 1941.[9]

The incidence of the reduction in consumption was unequal in different parts of the country, although it would be an intricate and unrewarding task to separate out the various elements. In some states the figures of consumption did not decline, because of war industrialization and the influx of population; in some, a similar result was produced by lax rationing.[1] What is certain is that the impact of gasoline rationing in the northeastern states was more severe than elsewhere, both because rationing was begun at an earlier date and because it was imposed with greater severity.

JAMES A. MAXWELL.
MARGARET N. BALCOM.

CLARK UNIVERSITY

9. Public Roads Administration, Classified Estimate of Travel by Motor Vehicles, 1936 to 1944. Table VM–1, 1945.
1. The Public Roads Administration figures show at one extreme that highway consumption in Utah, Nebraska and South Dakota was 85 per cent, 79 per cent and 79 per cent in 1943 compared to 1941, while in Connecticut, Massachusetts and Pennsylvania the figures were 57 per cent, 58 per cent and 59 per cent.

A NOTE ON MAXIMIZATION CRITERIA[1]

The purpose of this note is to re-examine some aspects of the maximization problem recently discussed by Lutz,[2] and to suggest that the quantity to be maximized depends upon circumstances surrounding an investment opportunity — particularly upon input limitations and the technological possibilities of timing input and output. For convenience, one of Lutz's examples will be used in this note. While he considered three alternative maximization criteria,[3] it is proposed to consider only two in the present note.

The two possibilities to be considered are the maximization of the internal rate of return — i.e. the rate at which costs and receipts must be discounted to give the investment opportunity a present value of zero,[4] and the maximization of the difference between discounted receipts and discounted costs, where both receipts and costs are discounted at the market rate of interest.[5]

In comparing these two alternative criteria, Lutz constructs the following example.[6] Suppose an entrepreneur is considering the investment opportunity presented by the growing of a tree, and that the costs[7] are all incurred at the time of planting. Assume also that the entrepreneur has perfect foresight and knows the

1. The author is indebted for criticism and suggestions to Dr. Gerhard Tintner, Dr. O. H. Brownlee and Dr. John Nordin.

2. F. A. Lutz, "The Criterion of Maximum Profits in the Theory of Investment," this JOURNAL, November, 1945.

3. Lutz also considers the possibility of maximizing the ratio between discounted receipts and discounted costs. Since this leads to the same result as maximizing discounted net receipts in the only example considered in the present note, it is excluded from consideration here.

4. K. E. Boulding, "The Theory of a Single Investment," this JOURNAL, May, 1935, pp. 475ff.

5. P. Samuelson, "Some Aspects of the Pure Theory of Capital," this JOURNAL, May, 1937, pp. 497ff. See also other authors cited by Lutz. For brevity this criterion is sometimes referred to as the maximization of discounted net receipts in the present discussion.

6. Lutz, op. cit. p .64. The present writer adopts the essentials of Lutz's example and the diagrammatic scheme. The notation is changed slightly to facilitate some of the mathematical demonstrations included in the present note.

7. It should be noted that Lutz includes rent paid for the use of the land as a cost, but at no point does he take account of the fact that the rent charge would be higher for a long production period than for a short one.

lumber value of the tree at all future dates at which he may cut the tree and sell the lumber, and that the value of the lumber increases at a decreasing rate.

The situation can then be represented by a diagram as in Figure I. The logarithm of the cost or value at a particular date is measured along the y-axis, time along the x-axis, with O the time of the initial investment. C represents the initial cost of planting the sapling. $R = f(t)$ represents the value of the lumber at all contemplated dates. The convenience of plotting the log of value quantities lies in the fact that if any quantity is cumulated at a constant rate over time, the resulting accumulated value appears as a straight line. Thus, different rates of return on the initial investment C can be represented as straight lines emanating from C, with the steepness of the slope indicating the rate of return. Obviously, the internal rate of return is maximized at the slope indicated by line II, if the producer sells the tree for K after lapse of time M. If line IV represents cumulation (or discount) at the market rate of interest, then V_2 represents the present value at time O of expected receipts L after lapse of time N when returns are discounted at the market rate. It is apparent that $V_2 - C$ is
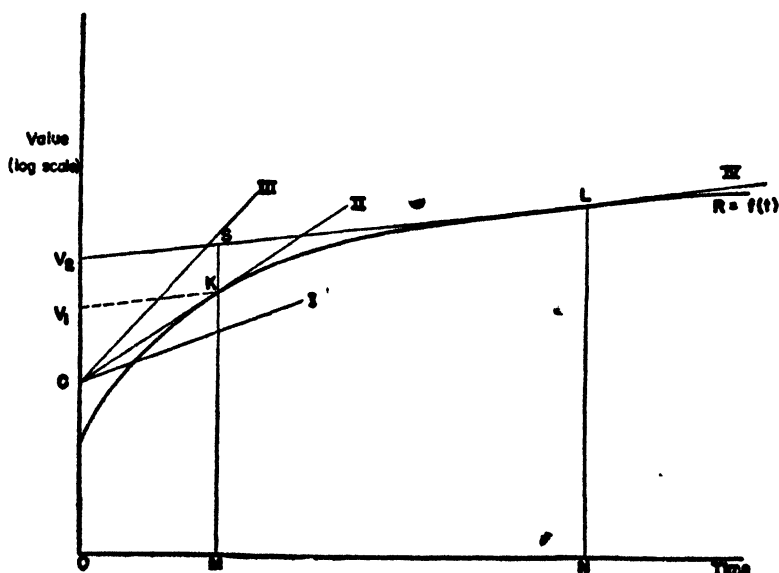
FIGURE I

the maximum difference between discounted receipts and discounted costs, where both are discounted at the market rate of interest. Thus maximization of the internal rate of return involves disinvestment at M, and maximization of discounted net receipts involves disinvestment at N.

In deciding between these two alternatives, Lutz writes, "Provided that a market rate of interest exists, and there are always people who value capital assets by capitalizing their yields at the market rate of interest, there can be no doubt that the entrepreneur will maximize $V-C$, not the internal rate. The value of the tree, not cut but standing, would move along the discount line] $[V_2L]$. At M its value standing would thus be $[S]$, which is more than K, its value as cut lumber. The entrepreneur would, therefore, under no circumstances cut the tree at $[M]$. Either he will hold on to it until $[N]$ or, if he wants to disinvest at M, he will sell it at the price $[S]$. The purchaser by letting it grow until $[N]$ will just make the market rate of interest on his investment."[8]

It should be noted that if it is possible to cut and sell and immediately reinvest at M, a preferable course of action is open. The entrepreneur could sell for K at time M, withdraw an amount equal to $K-C$, and reinvest an amount equal to C. After another lapse of time equal to M he again sells for an amount equal to K and reinvests C. This process is shown in Figure II. If for any reason the producer wants to disinvest at N, he has lumber worth k and three withdrawals worth $3(K-C) \cdot k+3(K-C) = W$ is greater than L in this case, and will always be greater if the entrepreneur can invest his withdrawals at the market rate of interest. In this case $W > L$, even if we assume that the entrepreneur cannot invest his withdrawals at all.

An even more favorable result occurs if we assume that the whole sale value at K can be reinvested at the maximum internal rate of return. This situation can easily be illustrated by assuming that the initial cost (C in Figure III) represents the cost of planting a number of saplings, and that at K_1 the whole amount is reinvested in a new and larger number of saplings. In Figure III we have the entrepreneur disinvesting at M and reinvesting in saplings worth B_2 at time of planting. $R_2 = f_2 (t-M)$ now represents the

8. Op. cit., p. 68.
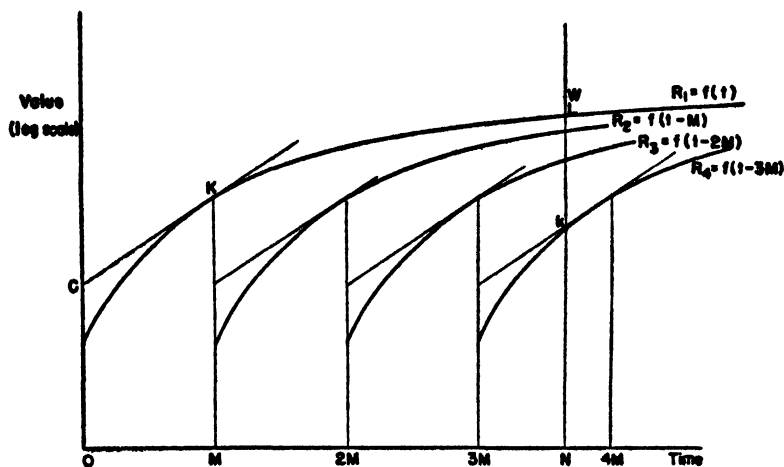
FIGURE II

growth function. It is apparent that $\dfrac{C_1}{B_1} = \dfrac{K_1}{B_2}$ and that the linear

distance B_2K_1 equals B_1C_1, since value is measured on a logarithmic scale. Another convenience of the logarithmic scale is that the
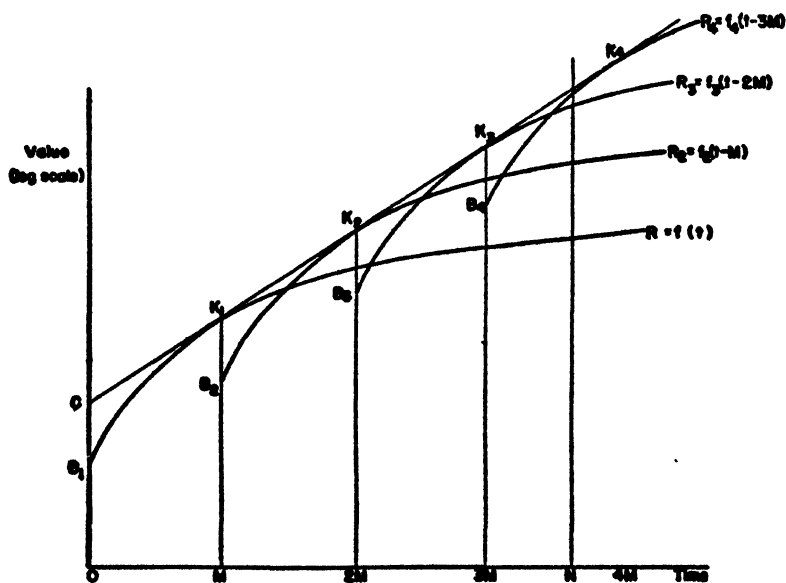
FIGURE III

new growth curve R_2 will have the same appearance as the original growth curve R.

The entrepreneur allows these trees to grow until their value is K_2, at which point he sells the lumber and reinvests in saplings with value B_3. At time $3M$ these have attained a value of K_3. This process would continue with the growth curve always becoming tangent to the line of maximum internal return until the entrepreneur desires a withdrawal or until the increased production begins to affect price. The value of the lumber available at N is W, which can be seen to be greater than L. The present value of W will then be greater than the present value of L at any rate of discount. The entrepreneur would probably not sell at N, but would choose $3M$ or $4M$ if a withdrawal were desired at about that time.

It is difficult to imagine the entrepreneur waiting until N to disinvest, if sale and reinvestment can be accomplished quickly and without too much cost at M. This might be possible in some lines of production and not possible in others. For example, a commercial poultry producer (particularly one in a southerly climate) can presumably sell off his flock and start new chicks at any time. He would thus sell at M and reinvest part or all of his funds at that time. On the other hand, prices of range steers for grain fattening are usually favorable only in the fall or early winter. Growth curves for cattle[9] suggest that both points K and L are reached before the start of a new production season. A beef-feeder would, therefore, presumably feed until the point L is reached, and then wait until the next season to reinvest.[1]

Another argument advanced by Lutz[2] in connection with this example supposes that the entrepreneur (refer again to Figure I) plants a tree at each moment in time. Once full production is reached, he both plants and cuts a tree at each moment in time. He plans to continue this operation indefinitely. Lutz then shows that if he cuts each tree at age M, his gross income stream, after full production is attained, is K per moment of time with a present value at time O of $\dfrac{Ke^{-rM}}{r}$, where r is the market rate of interest.

9. USDA Technical Bulletin No. 900, September, 1945, p. 9.
1. If he uses his own capital and has no other use or outside investment for it, he will better his position by waiting to disinvest until the value curve has a slope of zero, providing the value curve shows ~ale price less cumulated feed costs.
2. Op. cit., p. 69.

If he cuts each tree at age N, his gross stream becomes L per moment with a present value at time O of $\dfrac{Le^{-rN}}{r}$. In either case, the present value of the cost stream is $\dfrac{C}{r}$. Since Le^{-rN} is equal to V_2 in Figure I and Ke^{-rM} is equal to V_1, he concludes that the present value of the anticipated net income is greater if period of growth N is chosen.

This conclusion is correct under the assumptions Lutz makes. It should be observed, however, that the longer the period of production[3] chosen, the larger will be the grove of trees in existence after full production is reached. For example, if m represents a moment of time long enough to cut and replant a tree, there will be $\dfrac{N}{m}$ trees in existence under full production with period N, and $\dfrac{M}{m}$ trees if the period is M. If we take a grove of a given size, say $\dfrac{N}{m}$ trees, then the present value of the income stream where cutting occurs at M becomes $\dfrac{N}{M}\dfrac{Ke^{-rM}}{r}$ and the present value of the cost stream becomes $\dfrac{N}{M}\dfrac{C}{r}$. This is to be compared with discounted revenues of $\dfrac{Le^{-rN}}{r}$ and discounted costs of $\dfrac{C}{r}$ in Lutz's example above. Obviously, $\dfrac{N}{Mr}\,(Ke^{-rM}-C)$ will be greater than $\dfrac{1}{r}$ $(Le^{-rN}-C)$ if $\dfrac{N}{M}$ is greater than $\dfrac{V_2-C}{V_1-C}$. That this is true with the particular growth curve and discount rate shown in Figure I is apparent.[4] It might be useful, however, to consider the optimum period of production for this sort of investment opportunity under the general assumptions made by Lutz without restricting ourselves to the situation shown in Figure I.

Assume again a cost of planting C, a growth function $R=f(t)$, such that the first derivative $R'>0$, $R''<0$, $R(_o)<C$. Assume that

---

3. This is similar in some respects to period of production problems that have previously been treated. See the discussions indexed under "period of production" in Stigler, Production and Distribution Theories, New York, 1941.

4. That this is true can also be implied from a consideration of Figure II.

capital or land limitations limit the size of the grove to T trees. The entrepreneur is to plant a tree at each moment $\dfrac{T}{(p)}$ in time, so that T trees will be planted when the first one (the one planted at time O) is ready for cutting. If p represents the optimum growth period, then a tree will be planted each $\dfrac{T}{p}$ interval of time. Once full production is reached, a gross income of $R_{(p)}$ will be available each moment of time with a present value of $\dfrac{T}{p} \cdot \dfrac{R_{(p)}}{r}$ at time p, where r is the rate of interest or discount, and a present value of $\dfrac{T}{p} \dfrac{R_{(p)}}{r} e^{-rp}$ at time O. The cost stream will have a present value at time O o $\dfrac{T}{p} \dfrac{C}{r}$. The present value of discounted net receipts at time O is then

$$(1) \qquad W = \frac{T}{pr}(R_{(p)} e^{-rp} - C)$$

If we wish to choose a period p to maximize this value we get

$$(2) \qquad \frac{dW}{dp} = \frac{T}{p^2 r}(pR'(p)e^{-rp} - prR(p)e^{-rp} - R(p)e^{-rp} + C);$$

setting this equal to zero and simplifying we get

$$(3) \quad R'(p) = rR(p) + \frac{R(p) - Ce^{rp}}{p} \quad \text{as a condition of the optimum}$$

period. It can readily be shown that this is less than the period N in Figure I, as long as profit is possible at all. At period N the present value V_2 of the receipts from the growth of a single tree is maximized.

$$V_2 = R_{(N)}e^{-rN} - C$$

$$\frac{dV_2}{dN} = R'_{(N)}e^{-rN} - rR_{(N)}e^{-rN} = 0$$

$$(4) \qquad R'_{(N)} = rR_{(N)}$$

As long as profit is possible, $R_{(p)}$ must be greater than Ce^{rp} in equation (3). $\dfrac{R_{(p)} - Ce^{rp}}{p}$ will be positive; therefore R' will be greater if the present value of the whole investment opportunity is maximized. Since $R'' < 0$, p must be shorter than N and $R_{(p)}$ to the left of $R_{(N)}$. If $R_{(p)} = Ce^{rp}$ for the optimum period, then the market

rate of interest is equal to the maximum internal rate of return and the present value is maximized at O.

Another interpretation of equation (3) may be of interest. If $r = 0$, then

(5) $$R'_{(p)} = \frac{R_{(p)} - C}{p}.$$

This can readily be seen to be a condition of the period which would maximize the net income per unit time, once full production were attained. This period could be determined by drawing a straight line from C tangent to the receipts curve on a diagram which showed both time and value on linear scales.

While it would be dangerous to generalize too much from the simplified situations examined above, the writer feels that certain inferences can be drawn. The most important of these is that rational maximization criteria for individual investment opportunities are likely to depend on the technological possibilities and input limitations associated with each investment opportunity.

For example, if the period of production is technologically or externally determined (as in the beef-feeder example p. 160) and is longer than time N in Figure I, then the producer should not disinvest before N, and he should disinvest at N if we assume perfect mobility of capital at the market rate of interest (i.e. the entrepreneur can either borrow from the outside or make outside investments at the market rate). On the other hand, if the entrepreneur can disinvest and reinvest at will, as in the poultry example, i.e. if he can determine the period of production, then there are circumstances in which it will be rational for the entrepreneur to terminate each individual investment before the present value of net receipts from that individual investment has been maximized. This will be true even if we use the difference between receipts discounted at the market rate and costs discounted at the market rate as the criterion for evaluating the operation as a whole.

If initial capital is limited and the entrepreneur can forego withdrawals for several production periods, as illustrated in Figure III, the investment will grow fastest if the internal rate of return is maximized. In general, one might expect situations in which maximization of the internal rate of return is rational to be commonly associated with situations in which the capital input is limited.

In Lutz's example reproduced on p. 158, there is an implicit

limitation on the labor input — i.e. the rate of cutting and replanting is held constant regardless of the length of the production period. The size of the investment becomes a variable dependent on the production period chosen. In this circumstance, maximization of discounted net receipts on each individual investment has been shown to be the rational criterion.

If the entrepreneur wishes to maintain the operation at a given size (T trees) and wishes to maximize the discounted net receipts of the whole investment, then the period chosen is, in general, neither that which maximizes the internal rate of return nor that which would maximize the discounted net receipts on the growth of an individual tree (equation (3), p. 162).

This suggests to the writer that it might be profitable to construct various theoretical models involving various input and technological limitations likely to be encountered in practice, and to try for results with more general applicability than those of the example discussed above. Lutz correctly states that, "Under competitive conditions the situation depicted in the diagram (Figure I) cannot last. Either the profit to be obtained by investing in trees will attract other investors until, through a rise in·the costs of investment and a fall in the value of trees, the present value of the output becomes equal to the cost of investment, i.e., until (V_2-C) equals zero; or 'speculators' will start trees at the cost of (c) and sell them at once for (V_2) repeating the procedure until the difference . . . is eliminated. In this equilibrium situation the internal rate of interest becomes equal to the market interest rate. . . . "[5]

That is to say, the maximum internal rate of return will always be equal to the market rate of interest in long run equilibrium under perfect competition. Wherever there is a possible difference, it may be attributed to short-run considerations or market imperfections. These latter situations are worth considering and, as has been shown above, may lead to various rational maximization criteria.

<div align="right">CLIFFORD HILDRETH.</div>

IOWA STATE COLLEGE

5. *Op. cit.* p. 69.

SOME PRINCIPLES OF COMPENSATORY ACTION[1]

Some of the strongest opposition to, and greatest misunderstanding of, the Employment Act of 1946 during its legislative course stemmed from a fear of the way in which public monies might be spent. Some principles governing compensatory policy were vaguely formulated at different stages in the legislative process, but they appeared more as safeguards or considerations than as fundamental principles guiding public policy. The final Act does little to clarify the issue.

At least six major principles should guide compensatory action that is used as an instrument for maintaining maximum levels of employment. In approximate order of importance compensatory action should:

1. Give the greatest practicable encouragement to the free enterprise system;

2. Obtain the maximum amount of useful goods, services and/or assistance per dollar expended;

3. Create the maximum practicable number of additional man hours of useful employment directly and indirectly;

4. Maintain the normal distribution of the nation's labor force by occupation and by industry group, as nearly as possible;

5. Preserve the skills, work habits, and morale of individual workers to the maximum practicable degree; and

6. Contribute to an expanding exchange of goods and services among nations.

First of all, compensatory action must give the greatest practicable encouragement to the free enterprise system. This principle may, however, frequently run into direct and indirect conflict with the other basic principles, and is the crux of some of the sharpest disagreements on public policy.

On the one hand, efforts to place the unemployed on the most useful projects and to preserve as nearly as possible the normal distribution of the nation's labor force may push public activities into areas that have been traditionally regarded as private; for example, into housing and electrification. If this is done without regard to the effect on profit prospects, it may discourage as much

1. The writer wishes to acknowledge the helpful suggestions of John F. Timmons.

new private employment and investment as it initially creates. For if private investors and producers anticipate or fear constantly expanding government competition within their area, they are going to be increasingly reluctant to undertake new enterprise, and may even fail to maintain their past level of expenditures and employment. If public spending discourages business spending or investment, it will create just that much more need and justification for further expansion of public spending, thereby intensifying the vicious circle. The end result might be that an increase in public spending, purportedly for the preservation of the free enterprise system, in reality might undermine and destroy it.

On the other hand, if regard is given to the effect that unbalanced budgets or potential government competition are purported to have on private enterprise to the exclusion of all other considerations, then that may stymie governmental attempts to achieve maximum employment. Therefore, it will be exceedingly important that public work, which potentially competes with private enterprise, be evaluated in both its long-run and its overall effects, as well as in its short-run and immediate impact.

Some public work clearly displaces private enterprise and is to that extent a threat to the system. For example, extensive production and sale of furniture on the open market or the indiscriminate entrance of the Government into residential housing would be of this nature. It is possible, however, that a type of public work which may appear to be competitive in its immediate effects may encourage the private enterprise system in its long-run and overall effects. For example, the TVA, by rehabilitating an economically depressed region and making cheap electric power and waterway transportation available for private consumption and commercial use, has been a boon to private enterprise, not only in the region but throughout the nation as well. Or government entrance into low-income housing, which has been traditionally unprofitable for private builders, might create a demand for materials and equipment that would otherwise never arise.

Enough has been said to indicate that sound public spending policy must come to grips squarely with this issue. Compensatory tax policy must also be carefully weighed in terms of its effects on private enterprise.

Second, it almost goes without saying that the maximum amount of useful goods, services and/or assistance should be

obtained per dollar expended. No program of governmental expenditure will long be justified by public opinion unless the funds are spent not only for a good purpose but also for the best purpose at a given moment. One of the first tasks, therefore, is to place social priorities on different types of public spending. It will be difficult to apply this principle because of the lack of a common denominator with which to compare the social value of a dollar expended on building or improving schools, planting trees, developing recreational resources, increasing social security benefits, or subsidizing employment in private industry. Nevertheless, policy judgments have to be made and are, in fact, constantly being made upon the relative importance of these types of expenditures.

Third, compensatory action should be taken in such a way as to create the greatest practicable number of additional man-hours of useful employment, directly and indirectly. Application of this principle will call for discriminating economic research. It will be necessary to make continuous and careful analysis of the multiplier effects of government spending and the stimulating effects of lowering different taxes. It will be important to know whether the greatest economic impact per dollar of expenditure is to be obtained through public work, direct aids to consumption, or subsidizing private employment. Within the category of public work it will be necessary to know the amount of onsite, offsite, and tertiary employment created by each major type of work under different conditions. Moreover, the stimulating effects of spending policy must be equated with those of tax policy.

Fourth, compensatory action should maintain the normal distribution of the nation's labor force by occupation and by industry group, as nearly as possible. The normal distribution may be assumed to be the best, since it yields the goods and services that society wants most, as determined by free consumers' choices in the open market. It is also desirable to minimize the readjustments involved in shifting unemployed labor into temporary alternative employment and then back again into its normal place, once the economic maladjustment has been corrected.

Fifth, compensatory policy should preserve the skills, work habits and morale of individual workers to the maximum practicable degree, subject to the qualification that no skill should be preserved after it is rendered obsolete by technological change. It is too often overlooked that in the final analysis human values are

the most important. The economic wastes due to idle men and idle money during the 1930's were great, but the loss in skills, work habits and morale of millions of the nation's workers and their families was to a large degree irreparable. It has become clear, therefore, that compensatory action, particularly public work, should aim to maintain and improve the skills of the workers, so long as it is believed desirable to maintain them under the existing technology. It should encourage to the utmost honest and conscientious work habits. But most important is the preservation of the worker's morale by providing work of high social value at prevailing wages and technological efficiency.

Finally, all compensatory action should be tested by its effect on the world economy. It should look toward the promotion of world trade and peace, and should avoid measures or programs that lead to economic warfare and division among nations. Any other course invites a chain of events leading to consequences too horrible to contemplate.

E. Jay Howenstine, Jr.

Washington, D. C.

RECENT PUBLICATIONS

ABBOTT, CHARLES C. Management of the Federal Debt. New York: McGraw-Hill Book Company. 1946. pp. 194. $2.50.

ABRAHAM, E. W. A New Money for a New Age. Victoria, B. C.: The Author. 1946. pp. 8 [Copies may be obtained from the author (1125 Faithful Street).]

ANJARIA, J. U., LAKDAWALA, D. T., PANDIT, S. A. War and the Middle Class. Bombay, India: Popular Book Depot. 1946 pp. 33. 1–4 Rupees.

ATKINS, WILLARD, EDWARDS, GEORGE, MOULTON, HAROLD. The Regulation of the Security Markets. Washington, D. C.: The Brookings Institution. 1946. pp. 126. $2.00.

AUXIER, GEORGE W. Truck Production and Distribution Policies of the War Production Board and Predecessor Agencies, July 1940 to December 1944. Washington, D. C.: Civilian Production Administration. 1946. pp. 116.

BANQUE DES REGLEMENTS INTERNATIONAUX. Reglementation des Paiements avec L'Etranger. Bâle, Switzerland: The Bank. 1946. pp. 100.

BHATAWDEKAR, M. V. Our Sterling Balances. Bombay, India: The Popular Book Depot. 1945. pp. 34.

BIGHAM, TRUMAN C. Transportation: Principles and Problems. New York: McGraw-Hill Book Company. 1946. pp. 626. $5.00.

BLODGETT, RALPH. Principles of Economics. New York: Rinehart & Company. 1946. pp. 668. $4.00.

BOLLINGER, LYNN L., PASSEN, ALAN, McELFRESH, ROBERT E. Terminal Airport Financing and Management. Cambridge: Harvard University Press. 1946. pp. 385. $4.25.

BOLTE, CHARLES G. The New Veteran. New York: Penguin Books, Inc. 1946. pp. 182. 25 cents.

BRODIE, HENRY. Selective Price Control in Canada, World War II. New York New York University. 1946. pp. 28.

BURNS, ARTHUR F. Economic Research and the Keynesian Thinking of our Times. New York: National Bureau of Economic Research, Inc. 1946. pp. 69.

BURNS, ARTHUR F., AND MITCHELL, WESLEY C. Measuring Business Cycles. New York: National Bureau of Economic Research. 1946. pp. 560. $5.00.

CARLSON, VALDEMAR. An Introduction to Modern Economics. Philadelphia: The Blakiston Company. 1946. pp. 337. $3.50.

CENTRE NATIONAL D'INFORMATION ÉCONOMIQUE. Les Problèmes Théoriques et Pratiques de la Planification. Paris, France: Presses Universitaires de France. 1946. pp. 349.

CHESSA, FEDERICO. L'Economia e la Guerra. Torino, Italy: G. Giappichelli 1945. pp. 208.

CHRISTENSEN, AKSEL E. Dutch Trade to the Baltic about 1600. Copenhagen, Denmark: Einar Munksgaard. 1941. pp. 480. Cr. 35.

CIVILIAN PRODUCTION ADMINISTRATION. Hide and Leather Policies of the War Production Board and Predecessor Agencies, May 1940 to December 1943. Washington, D. C.: War Production Board. 1946. pp. 100.

COLEMAN, CHARLES H. Shipbuilding Activities of the National Defense Advisory Commission and Office of Production Management, July 1940 to December 1941. Washington, D. C.: Civilian Production Administration Bureau of Demobilization. 1946. pp. 148.

CONDOIDE, MIKHAIL V. Russian-American Trade. Columbus, Ohio: Ohio State University. 1946. pp. 160.

CONFERENCE ON RESEARCH IN INCOME AND WEALTH. Studies in Income and Wealth, Volume 8. New York: National Bureau of Economic Research. 1946. pp. 297. $3.00.

CONVERSE, P. D. Retail Trade Areas in Illinois. Urbana, Illinois: University of Illinois. 1946. pp. 31.

CRUM, W. L., AND SCHUMPETER, JOSEPH A. Rudimentary Mathematics for Economists and Statisticians. New York: McGraw-Hill Book Company. 1946. pp. 183. $2.50.

CURRIE, A. W. Economic Geography of Canada. New York: The Macmillan Company. 1946. pp. 455. $3.50.

DADE, EMIL B. Migration of Kansas Population. Lawrence, Kansas: University of Kansas. 1946. pp. 28.

DAMALAS, B. V. La Crise du Capitalisme et le Problème de L'Économie Dirigée. Paris: Presses Universitaires de France. 1946. pp. 323.

—— Essai Sur L'Evolution du Commerce International, Les Théories, Les Faits. Paris, France. Presses Universitaires de France. 1940. pp. 455.

DYMES, D. M. E. Synthesis in Education. Malvern, England: Le Play House Press. 1946. pp. 80. 5s.

ELSAS, M. J. Umriss einer Geschichte der Preise und Löhne in Deutschland, Band IIa. Leyden, Holland: A. W. Sijthoff's Uitgeversmaatschappij. 1940. pp. 649. $13.10 unbound; $14.70 bound.

FEI, HSIAO-TUNG. Peasant Life in China. New York: Oxford University Press. 1946. pp. 300. $3.50.

FELLNER, WILLIAM, HALEY, BERNARD F. Readings in the Theory of Income Distribution. Philadelphia, Pennsylvania: The Blakiston Company. 1946. pp. 718. $4.25. [Edited by a Committee of the American Economic Association.]

HANSEN, H. America's Role in the World Economy. New York: Penguin Books, Inc. 1946. pp. 170. 25 cents.

HARDY, CHARLES O., WILLIAMS, KENNETH B., AND ELLIS, HOWARD S. Prices, Wages, and Employment. Washington, D. C.: Postwar Economic Studies, Board of Governors of the Federal Reserve System. 1946. pp. 94.

JONES, DRUMMOND. The Role of the Office of Civilian Requirements in the Office of Production Management and War Production Board, January 1941 to November 1945. Washington, D. C.: Civilian Production Administration. 1946. pp. 351.

KAEMPFFERT, WALDEMAR. Should the Government Support Science? New York: Public Affairs Committee. 1946. pp. 32. 10 cents.

KRAGH, BORJE. Sweden's Monetary and Fiscal Policy before and after the Second World War. Stockholm, Sweden: Svenska Handelsbanken. 1946. pp. 20.

LANGE, OSCAR. The Problem of Franco Spain. New York: Polish Embassy. 1946. pp. 22.

LESCOFFIER, FRANCOIS. L'Économie de Grand Espace. Paris, France: Librairie Générale de Droit et de Jurisprudence. 1946. pp. 210.

LITTLE, K. L. The Relations of White People and Coloured People in Great Britain. Malvern, England: Le Play House Press. 1946. pp. 8. 6d.

LORIMER, FRANK. The Population of the Soviet Union: History and Prospects. Geneva, Switzerland: League of Nations. 1946. pp. 289.

LYNCH, DAVID. The Concentration of Economic Power. New York: Columbia University Press. 1946. pp. 423. $5.50.

McALEER, JAMES A. Chronology of the War Production Board and Predecessor Agencies, August 1939 to November 1945. Washington, D. C.: Civilian Production Administration Bureau of Demobilization. 1946. pp. 26.

McCAULEY, MARYCLAIRE. The Closing of the Gold Mines, August 1941 to March 1944. Washington, D. C.: Civilian Production Administration Bureau of Demobilization. 1946. pp. 77.

MANTOUX, ÉTIENNE. The Carthaginian Peace—or the Economic Consequences of Mr. Keynes. Oxford, England: Oxford University Press. 1946. pp. 210.

MASON, EDWARD S. Controlling World Trade. New York: McGraw-Hill Book Company. 1946. pp. 289. $2.50.

MENDERSHAUSEN, HORST. Changes in Income Distribution during the Great Depression. New York: National Bureau of Economic Research. 1946. pp. 173. $2.50.

MISES, LUDWIG VON, AND TUCKER, RUFUS S. Economic Planning. New York: Dynamic America, Inc. 1945. pp. 29.

MONTGOMERY, MARY, AND OLAWSON, MARION. History of Legislation and Policy Formation of the Central Valley Project. Berkeley, California: U. S. Department of Agriculture, Bureau of Agricultural Economics. 1946. pp. 276.

NIELSEN, A. E. Production Credit for Southern Cotton Growers. New York: Kings Crown Press. 1946. pp. 194. $2.50.

NOGARO, BERTRAND. Cours D'Économie Politique. Paris, France: Domat-Montchrestien. 1946. pp. 451.

—— La Monnaie et Les Systèmes Monétaires. Paris, France: Librairie Générale de Droit et de Jurisprudence. 1945. pp. 250.

OFFICE OF ECONOMIC STABILIZATION. Report of the President's Committee on the Cost of Living. Washington, D. C. · Department of Labor, Bureau of Labor Statistics. 1946. pp. 422.

PARKS, BROTHER E. PATRICK. The Roman Rhetorical Schools as a Preparation for the Courts under the Early Empire. Baltimore, Maryland: Johns Hopkins Press. 1945. pp. 122.

PATTERSON, WILLIAM F., AND HEDGES, MARION H. Educating for Industry. New York: Prentice-Hall. 1946. pp. 229. $2.50.

POLITI, S. M. Algumas Realidades Americanas Vistas Atraves de Uma Viagem. Sao Paulo, Brazil: Institute of International Education. 1944. pp. 23.

—— A Racionalizacao e o Bem-Estar Economico da Coletividade. Sao Paulo, Brazil: Department of Industrial Production. 1946. pp. 109.

—— A Tecnica do Preco de Custo da Producao. Sao Paulo, Brazil: Department of Industrial Production. 1946. pp. 55.

ST. GEORGE, MAXIMILIAN J., AND DENNIS, LAWRENCE. A Trial on Trial. Washington, D. C.: National Civil Rights Committee. 1945. pp. 503. $5.00.

SCHECHTMAN, JOSEPH B. European Population Transfers 1939–1945. New York: Oxford University Press. 1946. pp. 532. $5.00.

SECURITIES AND EXCHANGE COMMISSION Statistical Bulletin. Philadelphia: The Commission. 1946. pp. 23.

SHENOY, B. R. The Bombay Plan, A Review of its Financial Provisions. Bombay, India: Karnatak Publishing House. 1944. pp. 44.

SHEPHERD, GEOFFREY S. Marketing Farm Products. Ames, Iowa: The Iowa State College Press. 1946. pp. 445. $4.25.

SIMPSON, W. W. The Jewish Problem. Malvern, England: Le Play House Press. 1945. pp. 8. 6d.

SITTERSON, J. CARLYLE. Aircraft Production Policies under the National Defense Advisory Commission and Office of Production Management, May 1940 to December 1941. Washington, D. C.: Civilian Production Administration. Bureau of Demobilization. 1946. pp. 168.

SMITH, J. RUSSELL, AND PHILLIPS, M. OGDEN. Industrial and Commercial Geography. New York: Henry Holt and Company. 1946. pp. 978. $4.90.

SMITH, LEONARD J. Collective Bargaining. New York: Prentice-Hall, Inc. 1946. pp. 468. $5.00.

SOUTHARD, FRANK A., JR. Some European Currency and Exchange Experiences: 1943–1946. Princeton, N. J.: Department of Economics and Social Institutions. 1946. pp. 23.

STRANG, IDA GODON. Peace Plan. Chevy Chase, Maryland: The Author. 1946. pp. 21. 25 cents.

SUBRAMANIAN, S. India, Guide to Current Official Statistics. Delhi, India: Manager of Publications. 1943. pp. 144. 4 shillings.

TINBERGEN, J. International Economic Co-operation. New York: Elsevier Publishing Company, Inc. 1945. pp. 208. $4.00.

VAKIL, C. N., ANJARIA, J. J., LAKDAWALA, D. T. Price Control and Food Supply. Bombay, India: N. M. Tripathi & Co. 1943. pp. 120. 3 Rupees.

VILLEGAS, LUIS TORON, AND TORRES, ADRIAN ESTEVE. Estudio de los yacimientos ferriferos de Mexico. Mexico: Investigaciones Industriales del Banco de Mexico. 1945. pp. 200. [Yacimentos del Grupo del Norte, Estados de Chihuahua, Coahuila, Durango, Nuevo Léon y Zacatecas.]

WADIA, P. A., AND MERCHANT, K. T. The Bombay Plan, A Criticism. Bombay, India: The Popular Book Depot. 1946. pp. 60.

WILTSE, CHARLES M. Aluminum Policies of the War Production Board and Predecessor Agencies, May 1940 to November 1945. Washington, D. C.: Civilian Production Administration Bureau of Demobilization. 1946. pp. 348.

WOLMAN, LEO. The Conditions of Labor Peace. New York: American Iron and Steel Institute. 1946. pp. 12.

MULTIPLE-PLANT FIRMS, CARTELS,
AND IMPERFECT COMPETITION[1]

SUMMARY

I. Perfect monopoly: the short-run marginal cost curve, 174; the long-run cost curve, 179; intraplant vs. interplant economies and diseconomies, 181; the problem of optimum size, 184 — II. Mathematical analysis: geometric proof, 185; algebraic proof, 187. — III. Oligopoly: non-economic considerations, 192; the short-run marginal cost curve, 194; the long-run curve, 196; total cost, 196; long-run equilibrium, 197; elements of instability, 199. — IV. American experience: mixture of types, 200; decline of dominant firms, 202; the milk distributing industry, 203. — V. Conclusions, 205.

Although theories of imperfect competition have long since found a recognized place in textbooks, varying degrees of dissatisfaction with them have persisted due to their restrictive assumptions. To a very considerable extent this situation must continue until we have many more detailed studies of specific firms and industries and their methods of operation. The difficulty is, of course, the procuring of information. Even if many corporations overcame their reluctance to open their books to economists, those fundamental and intimate details of corporate policy formation which are not translatable into bookkeeping entries would still be lacking. To some extent, however, the empirical evidence already available to us provides a basis for significant improvement in the theory. Accordingly, I plan in this article (1) to extend the analysis

to deal with multiple-plant firms[2] and (2) to present a generalized oligopoly solution which can be employed as a frame of reference in empirical studies.

I

To illustrate the problem of multiple-plant firms, I shall first consider a case of perfect monopoly. For simplicity assume a linear demand curve remaining constant throughout the discussion, completely unspecialized factors of production (precluding any
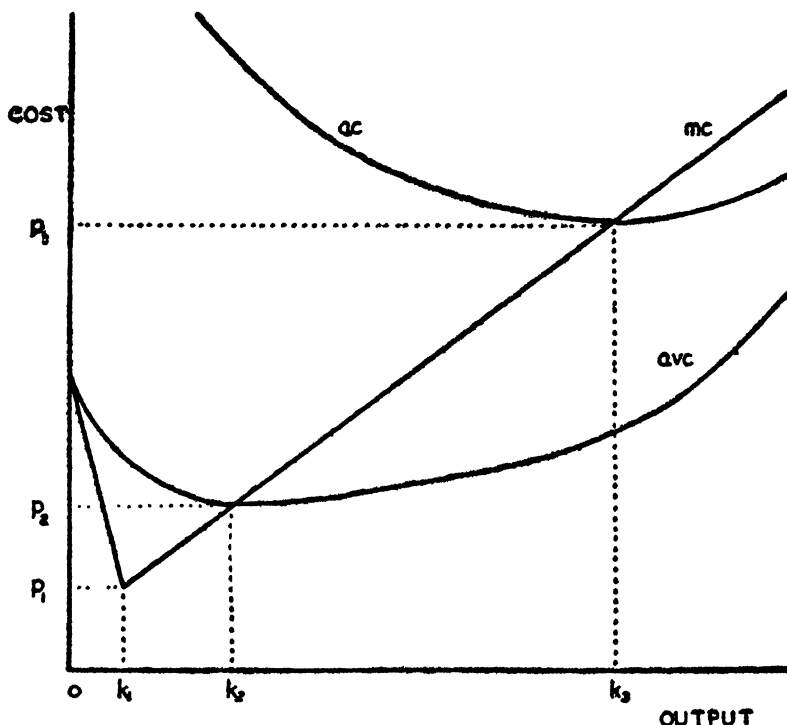
FIGURE I

monopsony power), and a single product. Assume that the monopoly firm consists of 100 individual (and identical) plants, each

2. The usual one-firm-one-plant analysis is completely unrealistic. Thus, in 1937 the 50 largest manufacturing corporations owned 2869 plants, or an average of 57.4 per firm. No firm owned less than seven plants, while one owned 497. (TNEC, Monograph No. 27, The Structure of Industry, pp. 675–714.)

of the long-run optimum size and with cost curves indicated in Figure I, where *ac*, *mc*, and *avc* represent average, marginal, and average variable costs, respectively. For convenience we have assumed that the plant marginal cost curve is composed of two linear sections.
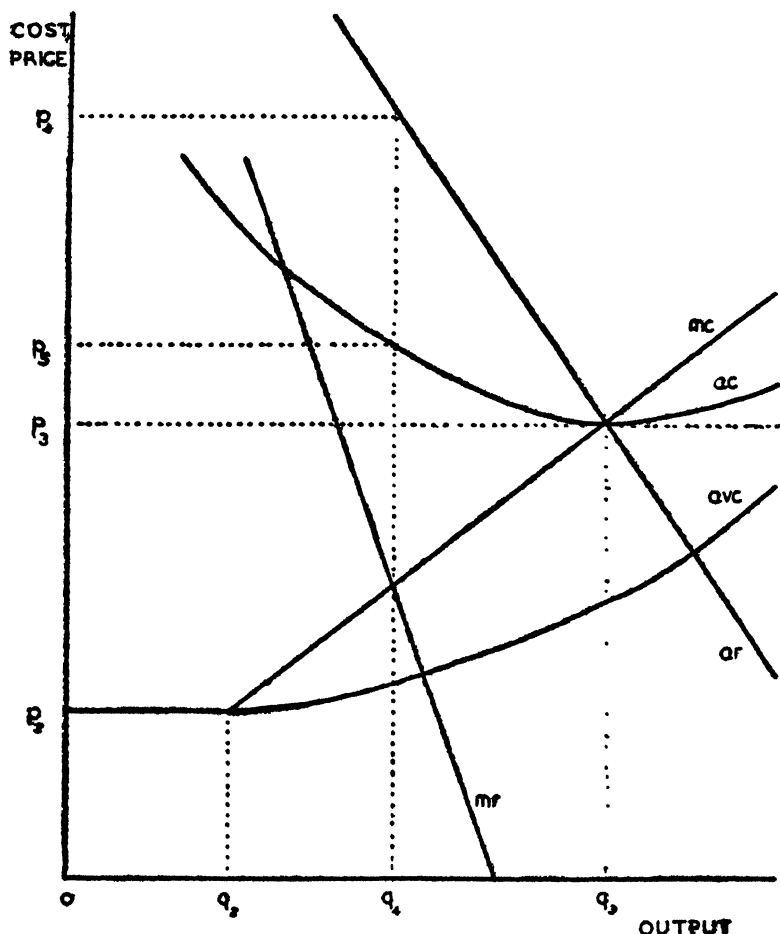
FIGURE II

We turn now to the problem of constructing the monopolist's short-run marginal cost curve (Figure II), noting that in the short-run (by definition) the monopolist cannot change the number of

plants in existence. Consider a given output q_i;[3] the question is, how will the monopolist allocate this output among the different plants in order to minimize costs? The answer falls into two parts, according as q_i is greater or less than $q_2 = 100k_2$. In the former case the optimum allocation would be to have the output equally distributed among the plants. This readily follows from Figure I. If one plant is producing more than another, its marginal cost will be higher; consequently, a reduction in total cost can.be effected by shifting the output from the former to the latter.[4] This process will continue until all plants are producing the same amount. We thus get the right-hand part of the marginal cost curve in Figure II. To construct the marginal cost curve for outputs less than q_2 we note that no plant *in operation* will produce less than k_2; any desired output $q_i < q_2$ will be produced by having x plants produce k_2 units apiece (where k_2 is the output corresponding to minimum average variable costs for the plant — cf. Figure I), with the remaining $100 - x$ plants left idle. (x is obviously determined by the relationship $xk_2 = q_i$.) This may be proved as follows. In the short run the monopolist must bear the fixed expenses of the 100 plants. Therefore he will minimize total expenses for any given output by minimizing total variable expenses. Consider now the given output $q_i < q_2$. If this output were equally allocated among all plants, each plant would produce $k_i < k_2$ at an average variable cost of $p_i > p_2$. Then the total variable costs would be $100k_ip_i$. However, the total variable costs, if x plants were to produce k_2 apiece, would be xk_2p_2. But $100k_ip_i > xk_2p_2$, since $100k_i = xk_2 = q_i$ and $p_i > p_2$. This is perfectly general and holds for any $q_i < q_2$ and $k_i < k_2$. In this range the monopolist would operate keeping some plants idle. The average variable cost would be $xk_2p_2/xk_2 = p_2$. Therefore, until q_2, the average variable cost curve is a horizontal line at a height p_2. By definition, the marginal cost curve, in this range, coincides with it.[5]

Strictly speaking the linear shape of the marginal cost curve is only an approximation which is approached as the number of

3. The units of the abscissa of Figure II are related to those of Figure I by the equation $k_j = q_j/100$ for any j.

4. This holds for the cost curves usually dealt with, but is not general. Cf. Section II below.

5. The position of the marginal cost curve in this range can also be established by noting that it can never be less than p_2, for the total expenses can always be reduced by at least p_2 per unit simply by closing down a plant.

plants increases. The actual shape is pictured in Figure III; this may be considered as a "blowup" of Figure II (note the break in the vertical axis). For convenience, however, we have considered a firm consisting of only seven plants, each of the type described in Figure I. v_i is the average variable cost of the firm, if i plants (neither more nor *less*) are used to produce a given output in the cheapest way possible. The v_i have a common origin, since one unit of output will always be produced by one plant producing one unit, regardless of the number of plants. As i increases, v_i tends to flatten out, since any given increase in output will increase the output per plant less the more plants there are. From our previous discussion we know that v_i will reach its minimum at an output of ik_2 and a height of p_2 ($i = 1, 2 \ldots 7$); that is, the minimum points are equidistant and at the same height.[6] The intersection of v_i and v_{i+1} indicates where it would be profitable to employ an additional plant. The heavy kinked[7] curve is thus the relevant average variable cost curve of the firm. It is tangent to the horizontal line (p_2 units high) at the outputs nk_2 ($n = 1, 2 \ldots 7$). Similarly, we construct a marginal cost curve mc_i for each v_i. For higher values of i, mc_i will approach closer to v_i since the latter tends to flatten out. It is possible that ranges may exist for which mc_i becomes a step function. This is especially true for the rising part, where the assumption that output will be equally allocated is more probable; mc_7 is drawn on this assumption. Then an output of $7k_2 + 1$ is produced, with six plants producing k_2 and one producing $k_2 + 1$; an output of $7k_2 + i$ ($i = 2, 3 \ldots 7$) is produced by having $7 - i$ plants produce k_2 apiece, and i produce $k_2 + 1$ — so that the marginal cost for all these outputs is the same. Thus, for outputs greater than $7k_2$, mc_7 is a rising step function, with steps seven units wide. On corresponding assumptions, in Figure II the marginal cost curve for outputs greater than q_2 is a step function with steps 100 units wide. As long as the width of

6. These results are given more generally with the aid of some very neat mathematics in M. F. W. Joseph, "A Discontinuous Cost Curve and the Tendency to Increasing Returns," *Economic Journal*, Vol. 43 (1933), pp. 390–393. However, the exposition there is impaired by the unwarranted assumption that any output is equally allocated among plants. Cf. Section II of this article.

7. That is, a *continuous* curve with *discontinuous derivatives* at certain points. Curves of this type have been erroneously referred to in the literature as "discontinuous."
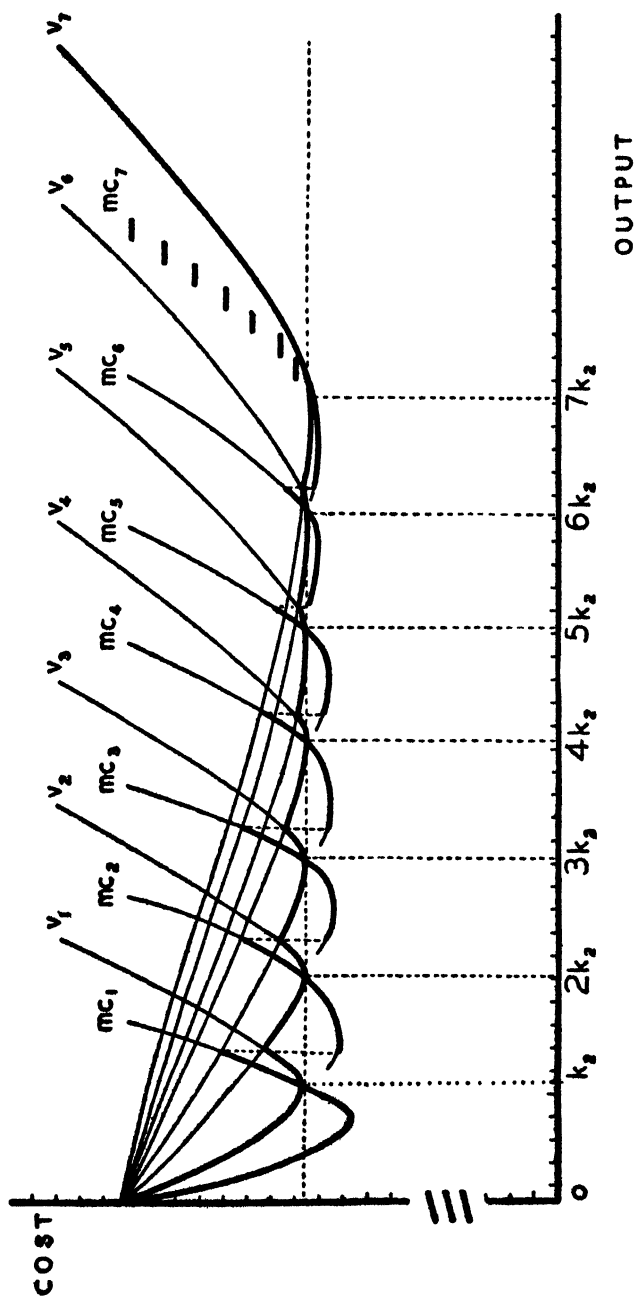
FIGURE III

the step is small relative to the scale of Figure II, we can approximate the rising part of the curve by a straight line.

In fact, due to the decreasing portions of the v_i and mc_i curves, we cannot construct these curves by allocating the output equally among the i plants. There do not seem to be any short-cut rules to follow, and the curves can be constructed only by trial and error allocations of the given output in different ways among the i plants and noting which way minimizes total variable cost. Similarly, it is impossible to determine by any simple rule when v_i and v_{i+1} will intersect — that is, when a new plant will be brought into operation. There might even be multiple intersection of v_i and v_{i+1}. For differently shaped plant cost curves we get entirely different results. Thus, for example, in the case of two plants we can construct cost curves such that the total variable cost is minimized by having one plant produce on the *rising* part of its marginal cost curve, and the other on the *falling* part (see Section II of this article).

Since the allocation of any given output is now determined, we can construct the other cost curves in Figure II, making use of the data in Figure I. For the demand curve ar the optimum output is at q_4, and the per-unit monopoly profit is $p_4 - p_5$. The usual textbook analysis stops at this point, with the implication that the monopolist is not producing at the point of minimum average cost (q_3) and should make no attempt to do so if he is to maximize profits. This certainly holds for the short run, but in the long run the monopolist can change his position by adjusting the number of plants through investment and disinvestment.

The problem is then one of determining the long-run cost curves of the monopolist (cf. Figure IV). The fundamental fact which must be noted here is that even in the long run the monopoly will not proceed to build different sized plants; investment and disinvestment in the firm will take place only by changing the number of plants. This follows from our assumption that the existing plants are of the long-run optimum size.

The construction of the monopolist's long-run average cost curve is analogous to the construction of the short-run average variable cost curve for outputs less than q_2. The process is identical if we note that in the long run (by definition) all costs are variable costs. We need only to observe that in the argument p_2 is replaced by p_3 and q_2 becomes infinitely large. In the long run the monop-
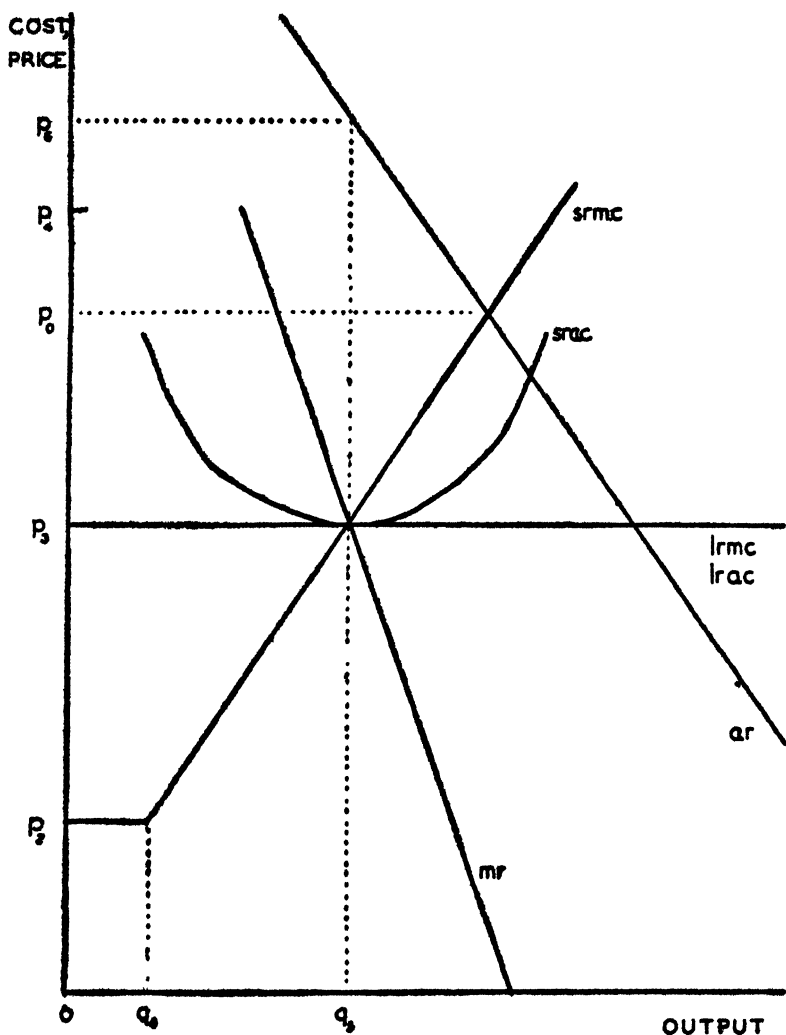
FIGURE IV

olist will minimize the cost of producing any output q_j by arranging his investment and disinvestment policies so that he will have exactly y plants, each producing k_s (cf. Figure I), where y is determined by the relation $y = q_j/k_s$. In other words, the optimum method of producing any given output is to have each plant producing at its minimum average cost point, and adjusting the

number of plants (by investment and disinvestment) so that the desired output can be produced. The total cost of *any* output q_j will then be yk_3p_3 and the average cost p_3, so that the long-run average cost curve (*lrac*) will be a horizontal line at the height p_3; the long-run marginal cost curve (*lrmc*) will coincide with it.[8]

The long-run equilibrium price and output in Figure IV (assuming *ar* to remain constant) are p_6 and q_6, respectively. The long-run price and monopoly profits are each greater than in the short run. Once the optimum long-run output q_6 is determined, the optimum number of plants in the long run — $m = q_6/k_3$ — is simultaneously determined. Thus, in the long run the monopolist will have m plants and the short-run average and marginal cost curves *srac* and *srmc* (Figure IV). *srac* will obviously have its minimum point at the output q_6, since for that output each of the m plants will be producing at its own minimum point k_3. Since $m < 100$, the marginal cost curve for the firm with m plants remains horizontal at p_2 over a shorter interval than when the firm consists of 100 plants. Specifically, for the case of m plants, *mc* remains constant only until $q_6 = mk_2 < q_2 = 100k_2$. It is interesting to note that in the long run the monopolist will *of necessity* be producing at the minimum point of his short-run average cost curve. We must now determine what particular assumption we have made that has led to this unusual result.

First let us distinguish between intraplant and interplant economies and diseconomies. Intraplant economies are what we usually have in mind when we speak of economies of large-scale production. These are derived from increases in the size of plant which enable use of more specialized and efficient machinery, develops skills in performing specialized tasks, eliminates movements of workers, and so on. Interplant economies are reductions in (social) cost following from the fact that two or more plants operate under a common management, instead of being separately owned. These take the form of economies in purchases and sales of materials, research, flexibility, "scientific management," risk-bearing, financing, integration, etc. Diseconomies of both types are due to increasing difficulties of coördination, bureaucratic inefficiency, etc. That there are substantial and continuing intra-

8. We must make reservations here analogous to those made above concerning the shape of the average variable cost curve for outputs less than q_2.
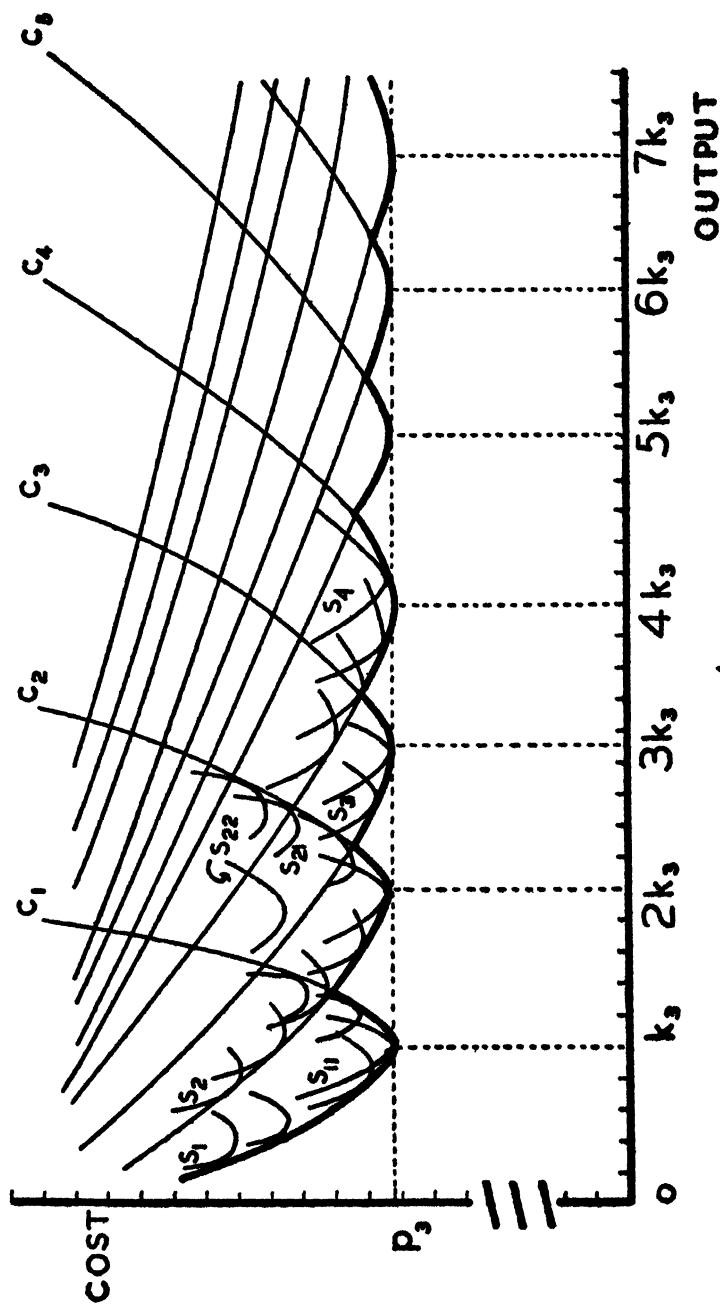
FIGURE V

plant economies has been shown to be true in many industries;[9] but on the question of interplant economies very little empirical evidence is available. Similarly, we have little information about the diseconomies.

Let us now consider the long-run average cost curve (Figure V).[1] The curve c_1 is the traditional Harrod-Viner envelope of the family of one-plant short-run cost curves s_1. The specific s_1 (say s_{11}) which is tangent to c_1 at the latter's minimum point is (by assumption) our curve of Figure I. Consider now any specific curve of the s_2 family — say s_{21}. This curve is constructed by allocating a given output in the best way possible among two plants of *arbitrary size*. We construct an s_2 curve for every possible combination of different size plants. We then construct the c_2 curve by marking off for any given output an ordinate equal to that of the lowest point on any member of the family of curves s_2 for that given output. Strictly speaking, c_2 is not an envelope curve, since there can obviously be members of s_2 which do not touch it: for example, s_{22} in Figure V. From our previous exposition we know that the s_2 curve tangent to c_2 at the latter's minimum point is that formed from two identical plants each of the size indicated in Figure I. A similar interpretation holds for the other s_{j}, and c_i curves. The latter will tend to flatten out as i increases. We know from our preceding discussion that c_i will reach its minimum for an output of $k_3 i$ at a height of p_3 $(i = 1, 2, \ldots \infty)$; that is, the minimum points are equidistant and at the same height.[2] The intersection of s_j and s_{j+1} indicates where it would be profitable for the monopolist to build a $(j+1)$-st plant. The heavy kinked curve is therefore the long-run average cost curve. In the case of other indivisibilities, the c_i curves would also be kinked, resulting in the long-run average cost curve having still more kinks than pictured here. The long-run marginal cost curve is discontinuous and of the same general shape of that in Figure III (for outputs less than $7k_1$), except that it extends indefinitely out, approaching more and more to the horizontal line at p_3. In order

9. J. M. Blair, "The Relation between Size and Efficiency of Business," *Review of Economic Statistics*, Vol. 24 pp. 125–135 (1942); Joseph Steindl, *Small and Big Business*, Oxford University Institute of Statistics, Monograph No. 1 (1945).

1. Figure V may be considered as a blowup of Figure IV; note the break in the vertical axis.

2. Cf. note 6 above

not to make Figure V too cumbersome, this curve has not been included.

The declining portion of c_1 measures the extent of intraplant economies: the greater k_3 the greater their importance. The importance of interplant economies is measured by the relative positions of the successive minimum points of the long-run average cost curve. There are four major possibilities. (1) They may lie on a horizontal line; this is the situation depicted in Figure V. It implies that there are no interplant economies or diseconomies, or that they offset each other identically for every output. (2) They may lie on a curve which remains a horizontal line for a significant distance and then begin to rise; here there are only interplant diseconomies. (3) The curve declines and then becomes horizontal; here there are only interplant economies. (4) The curve has the traditional U-shape pattern.

We see now that the previous unusual results for the monopoly case follow from our assuming situation (1) to hold; otherwise, there is no *necessity* for the monopolist to be operating at the minimum point of his short-run cost curve even in the long run. But it is essential to note that even in the traditional case (4), where from a purely probability viewpoint there is least likelihood that in the long run he will operate at the minimum point, the *probability* is greater than usually realized that his cost will be close to the minimum cost. This follows from the construction of the cost curves, which makes it impossible for the long-run *multiple*-plant cost curve (even if U-shaped) to be less flat-bottomed than the long-run *single*-plant cost curve.[3]

Formulation of the problem in this way focuses attention on the central policy problem of monopoly: optimum size of firm. Dissolution as the answer to monopoly is subject to two fundamental criticisms. (a) If firms are to be of optimum size, there might not be enough independent firms resulting from the dissolution to make the operation of competition possible. In other words, we will replace monopoly with some oligopoly situation, and it is quite possible that we would be as badly off as under monopoly. We shall deal more fully with this in Section III.

3. Though we have developed the theory of multiple-plant cost curves for monopoly only, it is clear that the construction of the cost curves is perfectly general and will hold in any of the cases of imperfect competition. These cases should be modified accordingly.

(b) Even if there are originally enough independent firms for competition to work, the situation might be unstable and develop into oligopoly.

Thus situation (1) is very favorable as far as (a) is concerned, depending only on the size of k_3 relative to market output. But the independent firms would then be in the familiar unstable situation of a long-run constant cost curve, and there would be no *economic* limit to their possible growth. The dissolution provisions would also have to define optimum firm and prevent firms from growing any larger.[4] Situation (2) is most favorable for a policy of dissolution, since it has the advantage over (1) that there is an economic limit to the growth of the independent firms. The case for dissolution is weakest under (3). Here again it depends on where the curve straightens out. Even if it is at a relatively small output (say $3k_3$, so that each competitive firm would consist of 3 plants), there would still be the problem of preventing indefinite growth. This problem would not be so bad under (4), but in either (3) or (4), if we insisted on making each plant a separate firm, we could do so only at the expense of efficiency in production. The cost would, of course, vary with the shape of the curve formed from the successive minimum points.

II

Our purpose in this section is to deal more mathematically with the short-run situation and show under what conditions equalization of marginal costs of the plants will not minimize

4. In the event that dissolution succeeded in placing each plant under separate ownership and making perfect competition work, in our preceding example we would have short-run equilibrium established at a price of p_0 (cf. Figure IV). This follows from our assumption that the ingredients are completely unspecialized, so that the supply curve for the competitive industry coincides with the marginal cost curve of the monopolist. This supply curve intersects the industry demand curve at the price p_0, thus establishing the short-run price for the competitive industry. Since this price exceeds minimum average cost (p_2) the industry is not in long-run equilibrium: firms will flow into the industry and the supply curve will shift over to the right until a price p_2 is established (again employing the assumption of a constant costs industry). This will be the case when there are 100 firms (with one plant each) in the industry (cf. Figure II). Thus, in Cassels' terminology, for the monopoly we have no long-run excess capacity (in situation (1)), though there is under-investment if we accept perfect competition as a criterion. (J. M. Cassels, "Excess Capacity and Monopolistic Competition," this JOURNAL, Vol. 51 (1936-37), pp. 440-443.)

costs to the firm.[5] We shall deal primarily with the case of a two-plant firm and offer (a) a geometric proof and then (b) a more general algebraic proof.

(a) Assume for convenience that each plant (A and B) has the same marginal cost curve reaching its minimum at b. For total outputs $x < 2b$ it is well known that equalization of marginal cost does not minimize cost (cf. (b) below). It is also obvious that in the case of a symmetric marginal cost curve an output of $2b$ can be produced either with each plant producing b or with one producing $b+d$ and the other $b-d$, where d is any positive constant less than or equal to b. We shall ignore these more trivial cases and consider $x = 2(b+1)$, which could be produced with each plant producing $b+1$. In order that it should be produced, instead, with A producing a and B producing $2(b+1)-a$, and *that this should be the only possible way of minimizing cost*, we have the following necessary and sufficient conditions (cf. Figure VI):

1. $A_1 < B_1$ — otherwise the entire output would be produced in B. This condition insures that *at least a* will be produced in A.

2. For any $0 < e \leq b+1-a$ it is true that $A_2 > B_2$. The most important value of $e = b+1-a$; for this insures that, even though we equalize the marginal costs of the two operating plants, we are not minimizing costs. The statement must hold for the other values of e to insure that no other point on the falling part of the marginal cost curve is the minimizing one.

These two conditions are very general and impose no restrictions inconsistent with the generally accepted U-shaped marginal cost curve. In general the class of cost curves meeting these conditions will have the following characteristics: (a) sharply falling initial stages, (b) flattening out for a short interval, then (c) rising even more sharply. (c) brings about condition (1); (a) and (b) together tend to bring about condition (2). To prove the first statement in this paragraph, we offer the following actual example.

Let A and B each have the cost curves described in Table I. We must now construct the v_1 and v_2 curves (cf. Figure III). The v_1 curve is, of course, represented by Table I. The v_2 curve is constructed in Table II. From a comparison of Tables I and II we see that v_2 intersects v_1 between outputs 11 and 12, so that the v_2 curve for an output of 12 is the relevant one. Yet this output,

5. This section may be omitted without disturbing the continuity between Sections I and III.

cost
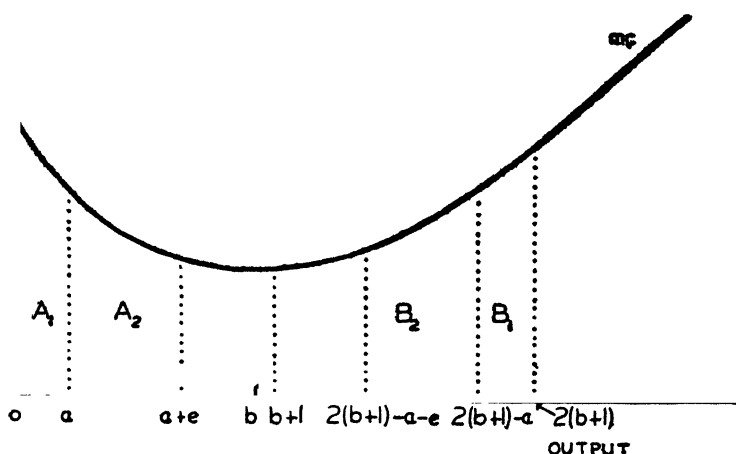
FIGURE VI

which could be produced by each plant producing at its minimum marginal cost output, is produced with one plant on the rising part of its marginal cost curve and the other on the falling.

(b) Consider again the case of a 2-plant firm.[6] Consider the plants as separate factors of production, and let $c(x_1)$ and $c(x_2)$ be the total variable cost curves of A and B, respectively, where x_1 is the output of A and x_2 of B. Then for any fixed output k the monopolist will seek to minimize

$$C = c(x_1) + c(x_2)$$

subject to the side condition

$$x_1 + x_2 = k$$

We employ the Lagrange multiplier and form

$$F = c(x_1) + c(x_2) - \lambda(x_1 + x_2 - k)$$

Minimizing F with respect to x_1 and then x_2 and eliminating λ we get as our first order conditions the familiar results

$$\partial c / \partial x_1 = \partial c / \partial x_2$$

In order that F should be a minimum, we need the second order conditions fulfilled[7]

6. I wish to express my appreciation to Trygve Haavelmo (University of Chicago) for his assistance in formulating the results of this paragraph.
7. Cf. J. R. Hicks, Value and Capital (1939), pp. 305ff.

$$D = \begin{vmatrix} 0 & \dfrac{\partial c}{\partial x_1} & \dfrac{\partial c}{\partial x_2} \\[2ex] \dfrac{\partial c}{\partial x_1} & \dfrac{\partial^2 c}{\partial x_1{}^2} & 0 \\[2ex] \dfrac{\partial c}{\partial x_2} & 0 & \dfrac{\partial^2 c-}{\partial x_2{}^2} \end{vmatrix} = -\left(\dfrac{\partial c}{\partial x_1}\right)^2\left(\dfrac{\partial^2 c}{\partial x_2{}^2}\right) - \left(\dfrac{\partial c}{\partial x_2}\right)^2\left(\dfrac{\partial^2 c}{\partial x_1{}^2}\right) < 0$$

TABLE I

Output	avc	tvc	mc
1	15.0	15	15
2	14.0	28	13
3	13.0	39	11
4	12.0	48	9
5	11.0	55	7
6	10.0	60	5
7	9.4	66	6
8	9.1	73	7
9	9.0	81	8
10	9.0	90	9
11	9.1	100	10
12	9.8	117	17.
13	11.1	144	27
14	13.1	184	40
15	15.9	239	55

TABLE II

| Output | Plant A | | Plant B | | TVC |
	output	tvc(A)	output	tvc(B)	tvc(A) + tvc(B)
1	1	15	0	0	15
2	1	15	1	15	30
3	2	28	1	15	43
4	3	39	1	15	54
5	4	48	1	15	63
6	5	55	1	15	70
7	6	60	1	15	75
8	7	66	1	15	81
9	8	73	1	15	88
10	9	81	1	15	96
11	10	90	1	15	105
12	11	100	1	15	115
13	7	66	6	60	126
14	7	66	7	66	132
15	8	73	7	66	139

Noting that $\left(\dfrac{\partial c}{\partial x_i}\right)^2$ is identically >0, and that $\dfrac{\partial^2 c}{\partial x_i{}^2} \gtrless 0$ $(i=1, 2)$ according as marginal cost is falling, at a minimum, or rising, we can formulate the following results:

1. For outputs $x=x_1+x_2<b$ (where b is the output for plant minimum marginal cost), each plant will of necessity be producing on the falling part of the curve. Therefore $D>0$ and equalizing marginal cost will *maximize* total costs. Minimum costs are achieved by having one plant produce the entire output.

2. If $x_1=x_2=b$, then $D=0$ and this allocation may be neither a minimum nor a maximum.

3. If $x=x_1+x_2>2b$ and we equalize marginal costs, then each plant is producing on the rising part of the marginal cost curve and $D<0$. This assures us that this allocation will minimize costs *relative to* all alternative allocations such that each plant is producing on the rising part of the marginal cost curve; that is, this allocation is the optimum one within a neighborhood such that every point is on the rising part of the curve. However, if we permit allocations such that one plant is producing on the falling part and one on its rising part, then for this allocation we may have $D>0$; thus the allocation achieved by equalizing marginal costs
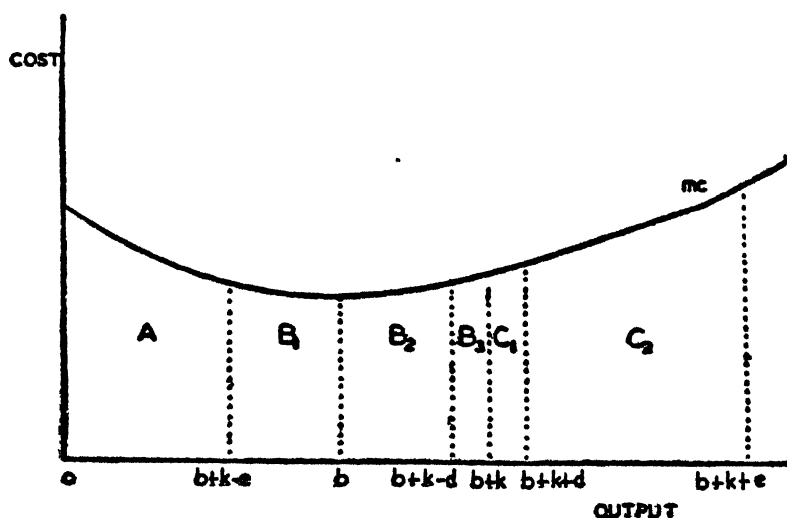
FIGURE VII

may be neither a minimum nor a maximum relative to the whole extent of the marginal cost curve.

These results can be represented graphically with the aid of Figure VII. From the shape of the marginal cost curve (which reaches its minimum at the output b) we know, for an output of $2(b+k)$, $C_1 > B_3$ as long as $d < k$. Therefore the total variable costs are less when the allocation is $b+k$ to each one than any other allocation $b+k+d$, $b+k-d$ (where $d < k$). Analogous results hold when total output is less than b: then from the shape of the curve we see that equalizing marginal cost will bring about a greater total variable cost than any other allocation. When, however, we permit the possibility of the allocation being $b+k-e$, $b+k+e$ (where $e > k$) we are no longer sure that $b+k$, $b+k$ will yield a smaller total variable cost than the former allocation. Specifically, the former allocation will be less when

$$2A + B + C < 2(A + B)$$

where $B = B_1 + B_2 + B_3$; $C = C_1 + C_2$.

This reduces to

$$C < B \text{ or } B_1 > C - (B_2 + B_3) > 0$$

Obviously by increasing the absolute value of the slope of the falling part of the marginal cost curve in the neighborhood of the minimum point, we can make B_1 as large as desired, while keeping the right side of this last inequality constant. Thus a marginal cost curve satisfying this inequality can easily be constructed.

These results can readily be generalized to the case of n plants, each with its own total variable cost curve $c_i(x_i)$, where the c_i are not necessarily the same. Our first and second order conditions are then:

$$\frac{\partial c_1}{\partial x_1} = \frac{\partial c_2}{\partial x_2} = \cdots = \frac{\partial c_n}{\partial x_n}$$

$$\begin{vmatrix} 0 & \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial c_2}{\partial x_2} \\[2ex] \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial^2 c_1}{\partial x_1{}^2} & 0 \\[2ex] \dfrac{\partial c_2}{\partial x_2} & 0 & \dfrac{\partial^2 c_2}{\partial x_2{}^2} \end{vmatrix} < 0 \qquad \begin{vmatrix} 0 & \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial c_2}{\partial x_2} & \dfrac{\partial c_3}{\partial x_3} \\[2ex] \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial^2 c_1}{\partial x_1{}^2} & 0 & 0 \\[2ex] \dfrac{\partial c_2}{\partial x_2} & 0 & \dfrac{\partial^2 c_2}{\partial x_2{}^2} & 0 \\[2ex] \dfrac{\partial c_3}{\partial x_3} & 0 & 0 & \dfrac{\partial^2 c_3}{\partial x_3{}^2} \end{vmatrix} < 0$$

$$\cdots \qquad \begin{vmatrix} 0 & \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial c_2}{\partial x_2} & \cdots & \dfrac{\partial c_n}{\partial x_n} \\[2ex] \dfrac{\partial c_1}{\partial x_1} & \dfrac{\partial^2 c_1}{\partial x_1{}^2} & 0 & \cdots & 0 \\[2ex] \dfrac{\partial c_2}{\partial x_2} & 0 & \dfrac{\partial^2 c_2}{\partial x_2{}^2} & \cdots & 0 \\[1ex] \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\[1ex] \dfrac{\partial c_n}{\partial x_n} & 0 & 0 & \cdots & \dfrac{\partial^2 c_n}{\partial x_n{}^2} \end{vmatrix} < 0$$

These results leave us with the conclusion that it is impossible to formulate any general rules to determine how many plants (in the short run) will be used to produce a given output, except the one proved in the section above, namely, with k plants, each with minimum average variable cost at $x_i = g$, outputs of ng ($n = 1, 2, \cdots k$) will be produced by having n plants each produce g. Furthermore, it is theoretically possible that $x = x_1$ will be produced with j plants, $x = x_1 + d$ with $j + 1$ plants, and $x = x_1 + d + e$ with j plants again. Consequently our Figure III should allow for the possibility of multiple intersection of the v_i curves.

Finally, we should note that these problems do not arise under

perfect competition, which may well explain why they have so long been neglected. In imperfect competition there is only a demand curve for the firm as a whole, and not for any individual plant. Therefore, before we can discuss equilibrium for the firm we must construct the aggregate cost curve for the firm as a whole. However, under perfect competition there exists a separate demand curve for each plant, namely, an infinitely elastic curve at the level of the market price. Consequently we can determine the equilibrium output of each plant independently of what takes place in other plants. In other words, the fact of the unlimited market which is present under perfect competition enables us to consider each plant separately.

Analytically this can be shown as follows. Consider a firm with n plants; the amount x_i is produced by the i-th plant, which has the total cost curve $g_i(x_i)$ $(i = 1, 2, \cdots n)$. Let $p =$ price of the product sold by the firm. By our assumption of perfect competition p is considered by the firm as given. Then the firm maximizes its profit

$$\pi = px - \sum_{i=1}^{n} g_i(x_i)$$

subject to

$$x = \sum_{i=1}^{n} x_i.$$

Substituting we have

$$\pi = p\Sigma x_i - \Sigma g_i(x_i)$$

from which follow our familiar maximizing conditions

$$\partial \pi / \partial x_i = p - \frac{\partial g_i(x_i)}{\partial x_i} = 0 \ (i = 1, 2, \cdots n)$$

that is, the marginal cost of each plant must equal the market price, our usual condition for equilibrium under perfect competition.

III

In this section I shall assume that the monopoly has been dissolved and replaced by one hundred independent firms, and then consider the arrangements that might grow up between them

in the absence of perfect competition. This failure of competition to develop may be due either to active desire to achieve monopoly gains, or to passive acceptance of noncompetitive arrangements due to interdependence and indeterminacy which make it impossible to adopt the rules of perfect competition.

We must recognize at the outset that in any realistic approach to the problem of monopoly and oligopoly we cannot deal in purely economic terms, but must introduce concepts and motivations which more closely approximate international power politics. On the one hand, corporation leaders have "corporationistic" feelings, together with a desire for power that is inherent in large size. On the other, corporations frequently undertake expansion programs for defensive purposes as well as for aggressive: vertical integration must be undertaken to assure strategic raw materials and market outlets— the corporation cannot allow itself to become dependent on other firms for these essentials. (The analogy to protectionism and war is complete.) In democratic societies the freedom of individuals becomes the ultimate limit to this integrative process; these societies prevent the corporation from achieving complete security by restricting its control over the factor of production labor, with its right to strike, and freedom of contract. Similarly, horizontal integration must be adopted, if the firm is to retain its position in the industry. The firm must accumulate large reserves, for in the event of a price war, victory is not to the most efficient but to the one with the largest reserves. In each case, it is the fear of imperfect competition which makes the corporation adopt methods of imperfect competition itself. This is what makes the oligopoly problem so difficult: it cannot be solved piecemeal. This vicious circle will continue until economists provide rules for the social control of oligopolies that will both protect the public and be workable: indeterminacy must be removed without leaving the door open to collusive exploitation. In brief, some criteria must be provided to distinguish "good" imperfect competition from "bad." It is quite likely that the controls devised will involve a much greater degree of direct government intervention than we have known heretofore.

Despite these qualifications, I now proceed to examine the workings of a market-sharing oligopoly arrangement among the newly independent firms. I shall attempt to show that this cartel

arrangement (as we shall refer to it) should be used as a general model for practical studies of imperfect competition. This is not to say that as it stands it is realistic; in fact, its assumptions will prove to be quite arbitrary. However, it does provide a convenient "jumping-off" point from which modifications can be made to deal with actual cases.

Assume that the one hundred independent firms set up a central office which decides on a common price and output policy for the industry. The cartel allocates quotas among the different firms in such a way as to minimize the costs of any given output. We no longer continue with our unrealistic (monopoly) assumption that the cartel can control entry into the industry. In fact, and this is the distinguishing feature, we assume that, as a result of anti-monopoly laws or the pressure of public opinion, there is free entry. Specifically, we assume that the anti-monopoly laws prevent single ownership of the industry and restrictions on entry, but permit agreements (either tacit or explicit) among the supposedly competitive firms. The cartel must thus permit any firm which wishes to do so to enter the industry and become a member of the cartel. In order to determine the short-run and long-run equilibria of the cartel, we must first construct its cost curves.[8] .

Let us first consider the short-run marginal cost curve. For outputs greater than q_2 (Figures VIII and II) the cartel will minimize costs by equalizing marginal costs among all firms and having each produce an equal amount. For outputs less than q_2 the cartel will follow exactly the same procedure as the monopoly in allocating production: x firms will produce k_2 units apiece, where x is determined by the equation $xq_2 = q_i$, where q_i is any output less than or equal to q_2. The short-run cartel situation is thus identical with short-run monopoly (Figure II). Each firm will produce k_4 to give a total output for the industry of q_4 and a price of p_4. The firms will share according to their quotas (and therefore equally) the cartel profits, which are equal to the area of the rectangle $p_4 r_1 s_1 p_{12}$.

In the long run the existence of these cartel profits will

8. The reader should make modifications to the shape and construction of these curves analogous to those pointed out for the curves in the preceding section. Since I have already dealt at length with this problem, it will be omitted here.
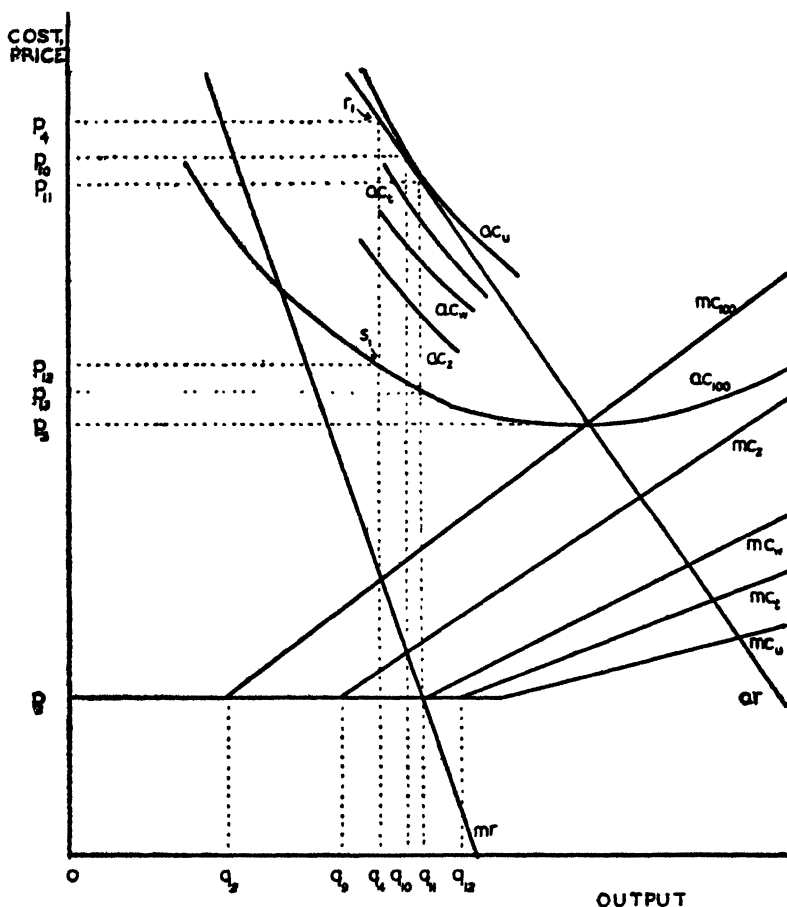
FIGURE VIII

attract new firms into the industry[9]. Assume that these too become part of the cartel. We must now consider what happens to the short-run cost curves of the cartel as new firms enter. By

9. New investment might of course also come from the old firms' expanding in order to increase their quotas and relative standing in the industry. Or they might wait until a new firm establishes itself, and then buy it up along with its quota. This last was the pattern of the German potash cartel and is also characteristic of the American meat packing industry. Cf. George W. Stocking, The Potash Industry (1931); Wm. H. Nicholls, "Market-Sharing in the Packing Industry," Journal of Farm Economics, Vol. 22 (1940), pp. 225-240.

assumption the existing plant is the optimum one; therefore, assuming no changes in technology or prices of factors, the new firms will build plants of exactly the same size. Assume also that one firm continues to operate only one plant.

As in the case of short-run monopoly, marginal cost can never fall below p_2. But in the long run, as new firms enter, the point corresponding to q_2 in Figure II moves over further to the right (cf. Figure VIII). Specifically, for any number of plants $z > 100$, the marginal cost curve (mc_z) will be a horizontal line at the level p_2 until the point $q_9 = zk_2$. Outputs up to this point will be produced by keeping some plants idle and the remaining plants each producing k_2. For outputs greater than q_9 (with z firms), the cartel will allocate quotas equally among the firms and the marginal cost curve will rise. The marginal cost curve for z firms not only is horizontal for a longer stretch than that for 100 firms, but the slope of its rising part is smaller; therefore it will always lie below the marginal cost curve for 100 firms. This is true because for any given increase in output the increase in marginal cost for z firms is less than that for 100 firms, since the increased output can be shared ₜₙong a greater number of firms. For example, for a given increase in the cartel output, with only 100 firms each one might have to increase output from k_1 to k_3; while for z firms, each one might only have to increase from k_1 to k_2, with a corresponding smaller increase in marginal costs (cf. Figure I). Another way of looking at this is to note that the "steps" in the rising part of the curve become wider as more firms enter (cf. above, p. 177).

So much for the marginal cost curve. As more firms enter, the total fixed cost, and therefore the average fixed cost curve, will rise uniformly. But the average variable cost for the cartel with any number of plants $z > 100$ will be less than, or equal to, the average variable costs with 100 firms. Until q_2 the curves will coincide as horizontal straight lines at the level p_2. For outputs from q_2 to q_9 the z curve will continue horizontally, while the other curve will begin to rise. For outputs greater than q_9 the average variable cost for z firms (avc_z) will also begin to rise, always remaining, of course, below its marginal cost curve. It will also lie below avc_{100}, since for any output $q_i > q_9$ each firm will produce a smaller output and will thus have lower average variable costs (cf. Figure I). Consequently there is no definite relationship between the average cost curve for z firms and that for 100 firms:

it might be higher in some intervals and lower in others. For example, if the cartel output were such that with 100 firms each firm were producing k_3 or less (Figure I), an increasing number of firms would reduce (for the fixed cartel output) the output per firm and drive each firm to the left and higher on its average cost curve. Therefore, the cartel average cost for this output would be greater with z firms than with 100. If, on the other hand, cartel output were such that with 100 firms each one were producing immediately to the right of k_3, then for a slight increase in the number of firms the cartel average cost for that output would be decreased, as each plant was pushed down to its minimum point k_3; while for larger increases in the number of plants each one would be pushed up on the falling part of the average cost curve until the cartel average cost was higher than with 100 firms. If, finally, with 100 firms the output is such that each plant is far to the right on its average cost curve, then even for large increases in the number of firms, cartel average cost would be reduced. However, for the output q_3 with 100 firms, each firm is producing at its minimum average cost. We have thus the first situation described here, and therefore for any given output $q_j < q_3$ the cartel average cost with a greater number of firms will be greater than with a smaller number. For outputs greater than q_3 the other two situations will hold.

From our previous discussion we see that, as new firms enter, the marginal cost curve is pushed uniformly to the right. Since we have assumed that the average revenue (and therefore the marginal revenue) curve remains constant, this means that in the long run cartel price will fall and output increase. Thus, for z firms we have equilibrium with output $q_{10} > q_4$ and price $p_{10} < p_4$.

Let us suppose that even with z firms there are still cartel profits; then new firms will continue to enter. Assume that the number of firms increases to w where $w = q_{11}/k_2$. The same relationships hold between the cost curves of the cartel with w firms and the cartel with z firms as between z firms and 100 firms. The equilibrium output will be $q_{11} > q_{10}$ and the price $p_{11} < p_{10}$. Assume, now, that even with w firms profits are still being made. Let the new number of firms be t, where $t = q_{12}/k_3$. Let us examine the effect of this new inflow of firms on the equilibrium situation.

The marginal cost curve (mc_t) is changed as indicated in Figure VIII. The first significant point is that *the equilibrium output has not changed* and is still at q_{11}. Furthermore, the output

q_{11} will be produced in exactly the same way as with w firms: w plants will produce k_2 each to yield a total output of $k_2 w = q_{11}$. The remaining $t - w$ plants will remain idle and be paid their fixed costs and aliquot share of the cartel profits.[1] The other significant point is that average costs for an output q_{11} with t firms *will definitely be greater* than for w firms. This follows because average variable cost is the same in both cases $(= p_2)$, while average fixed cost is greater in the former due to the additional fixed costs of the $t - w$ new firms. Thus per-unit profit is definitely smaller. If there are still profits, new firms will continue to enter, driving up the average cost curve for the output q_{11} until it is tangent to the demand curve at that output. At this point price will equal average expense and there will be no profits and no further inducements to enter. The industry will then be in long-run equilibrium.[2] It can be shown, however, that both in theory and in practice this equilibrium is a very unstable one.

The long-run equilibrium number of firms in the industry, u, can be determined as follows. Total profits P, when there are w firms, are

$$P = p_{11} q_{11} - (q_{11} p_2 + wf)$$

where f is the total fixed cost per plant and p_2 the average variable cost. Then

$$u = w + \frac{P}{f}.$$

That is, the number of new firms above w is limited by the amount of cartel profits available to pay them their fixed expenses in order to keep them idle.[3]

The long-run equilibrium is thus one of both excess capacity (in the sense that the cartel is operating below its long-run minimum

1. Theoretically, at this point a new firm could obtain its share of cartel profits by merely *threatening* to build a new plant.
2. The classic case of a cartel in such a long-run equilibrium is the German potash cartel, which in 1928 operated only 60 of 229 plants. B. R. Wallace and L. R. Edminster, International Control of Raw Materials (1930), Chap. 4.
3. It is interesting to note that the graphic equilibrium obtained here is similar to that of the familiar Chamberlin product differentiation case. But here the curves refer to the *industry* as a whole, not to the individual firm. Also, here the equilibrium is obtained solely by shifts in the cost curves, while there the main shift takes place through the demand curve (for the individual firm).

cost point) and overinvestment.[4] If we measure excess capacity in terms of the output that could be yielded if output were at a point where marginal cost equals price, the results are equally impressive. Although there are more plants in the cartel than under long-run competitive equilibrium, the industry output is less and the cost and price higher, with a very low (normal) rate of profit.[5]

Such long-run equilibria are highly unstable. This is clearly shown in our model by the discrepancy between the marginal revenue of the *firm* (which approximately equals price on the assumption of non-retaliation) and its own very low marginal cost. Each individual firm realizes the ease and profit with which it could sell additional units beyond its quota. Thus the temptation to "bootlegging," smuggling, and "chiseling" is strong. As cartel profits decrease with the influx of new firms, this pressure becomes irresistible, especially for the low-cost firms, and the eventual breakdown of cartel discipline is inevitable. The pernicious (for the cartel) fact remains that it is to the *maximum advantage* of each firm to stay out of the cartel and sell in unlimited quantities at the cartel price (or just below it), while all other firms remain members of the cartel and by their common restrictive policies hold up the price. This has been the pattern of breakdown of many cartels, with rubber (the Stevenson Plan) as the classic example.[6]

4. Whereas under monopoly there was no excess capacity and underinvestment. This last pernicious and wasteful effect of the cartel is what makes many economists believe that a situation of out-and-out monopoly is preferable to a cartel. Cf. above p. 185, note 4.

5. Excess capacity is here presented as the outcome of cartel operations, and the cartel itself is depicted as beginning from a situation of perfect competition. As is well known, however, excess capacity frequently first arises through shifts in demand or technological changes, precipitates a disastrous period of "cutthroat competition," which is finally ended by setting up a cartel arrangment. This pattern has been especially important among products with low income elasticities of demand — the so-called primary products (e.g. the rubber, coffee, and wheat cartels). Thus excess capacity is itself a *cause* of the cartel. Cf. J. W. F. Rowe, Markets and Men (1936); Wallace and Edminster, op. cit., W. Y. Elliott et al., International Control in the Non-Ferrous Metals (1937).

6. Cf. Rowe, op. cit.; K. E. Knorr, World Rubber and Its Regulation (1945); Rowe, "Studies in the Artificial Control of Raw Material Supplies: No. 2, Rubber," London and Cambridge Economic Service, Special Memorandum No. 34 (1931); C. R. Whittlesey, Governmental Control of Crude Rubber (1931). The last two are excellent critical studies of the Stevenson Plan.

That the same pattern was at work in the U. S. copper export cartel after the first World War is evident from the following testimony of C. F.

Another cause of instability lies in a fact from which our model abstracts by its assumption of uniform cost curves for all firms. There is a fundamental conflict of interest (within the cartel) between the low-cost and high-cost firms, with the latter insisting on high enough prices to cover their costs as the condition of their remaining in the cartel. There is also the very difficult problem of allocating quotas among the firms, which always creates much dissension and bickering. The forces described in this and the preceding paragraph go a long way in explaining the breakdown of many of our cartels. The cartel is in the unenviable position of having to satisfy everyone, for one dissatisfied producer can bring about the feared price competition and the disintegration of the cartel. Thus the successful cartel must follow a policy of continuous compromise.

In view of the difficulties of maintaining cartel discipline, it is not surprising that successful cartels have resorted to one or more of the following practices: (*a*) invoked government aid to compel membership and enforce cartel decisions (quotas and prices) — especially true of Europe; (*b*) controlled entry into, and operation within, the industry through patents — a frequent practice in the chemical industries; (*c*) controlled entry into, and operation within, the industry by ownership over the scarce raw material cartelized, e.g. tin, potash, lead; (*d*) compelled membership or prevented insubordination by dumping at (temporarily) greatly reduced prices in the market area of the non-coöperating producer. The cartel is frequently prevented from following this last practice by force of law (especially anti-dumping tariffs) or public opinion.

IV

In the United States open cartel arrangements of the type analyzed here are not frequent, since they are strongly discouraged by antitrust law — even more so than outright merger. Nevertheless, I shall show that many of the (tacit) arrangements which

Kelley (president of Anaconda Copper Mining Co.) before the TNEC (Hearings, Vol. 25, pp. 13164–13165):

"The Copper Export Association finally broke up due to two causes. One was the withdrawal of certain members, led by the Miami Copper Co. . . . There was an increase in competition from nonmembers abroad. There was a constant undercutting of price, and certain members felt *that they were holding the umbrella*, and it was more desirable to have freedom, and so gradually by withdrawals it lost its importance." (Italics mine).

do evolve in our economy have striking similarities to our cartel model. An unfortunate result of the classification of imperfect competition into several types is the failure to recognize that in actual life these types are inextricably mixed. Insofar as our economy can be characterized by a single pattern, I think it is one in which the given industry produces differentiated products and consists of a few (say three or four) very large firms doing the bulk of the business, plus many smaller "independents." The large firms act more or less as leaders for the industry in setting price policy, and so on. Some form of tacit or explicit market sharing arrangement (by percentages, market areas, recognized customers, etc.) exists to modify (if not remove) competition between the large firms. They might also proceed on the assumption that the other dominant firms will follow their prices both upwards and downwards. This gives results identical with our cartel model. The industry also has a trade association to help in maintaining discipline and implementing the price policies of the leading dominant firms.

The steel, petroleum, agricultural implements, anthracite coal, light bulb, cigarette, meat packing, and many other industries all fall within this general pattern.[7] In all these industries there has also been a decided tendency for the (original) dominant firms to decline in relative importance over the years. In some cases, this has been serious (steel, petroleum, meat packing); in others, relatively mild (anthracite coal, light bulb, cigarette). The decline has not taken place in absolute size; rather, the several industries have grown, but the dominant firms have grown at a slower rate. The dominant firms have apparently also attempted to pursue a policy of price stabilization, and have succeeded in varying degrees.

If we now interpret these facts in terms of our cartel model, we get very fruitful results. We must first consider the dominant firms as taking the place of the cartel "central office" and setting policies for the whole industry. They will discourage price-cutting by exhortation, "social" pressure, repeated stressing of the disaster which faces the industry as a result of price-cutting, and threats (explicit or implied) of underselling non-coöperating firms in their markets, if they persist. United States Steel and Standard Oil were, in their early years, notorious examples of this last practice.

7. Cf. A. R. Burns, The Decline of Competition (1936), especially Chap. 3 and pp. 140–145.

Through these methods the dominant firms are more or less able to maintain discipline within the industry and agreement on a common price. Even if the dominant firms are low-cost firms, they still may set a higher price than they themselves would prefer, in order to satisfy the other (high-cost) producers and prevent them from price cutting. As a rule, the dominant firms try to follow a policy of price stabilization. This may be due to the simplicity of the rule, or it may reflect the fear that arises every time the price is changed: whether the lead will be followed. In brief, the relationship between leaders and followers may be so delicate that the leaders take every care to prevent subjecting it to stress.

The decline in the relative position of the leader is readily explained as the familiar cartel phenomenon of the inflow of new firms and the expansion of old ones. The dominant firms themselves may not expand at the same rate as the industry, since they tend to be near their maximum size, and further expansion might involve them in many of the inefficiencies of large-scale operation. It is also possible that the dominant firms are high-cost producers and that the price set by them, though yielding relatively small profits in their case, would enable the other (low-cost) firms to earn much higher profits, thus increasing their incentive to enter the industry and expand. Finally, it should be noted that the decline in relative position may be due to weakness in the control exercised by the dominant firms. During periods of declining demand the "independents" will indulge in much more price cutting in order to increase their sales. It is quite possible (and seems to have been the case in steel and copper, for example) that the dominant firm will continue with its stabilized higher price and not retaliate for a time; that is, it expects it will still profit by this policy, although it is in the position of "holding the umbrella" for the other firms and restricting its own output relatively more than theirs.[8]

Finally, it is very instructive to examine those industries in which the dominant firms have declined relatively little. The results are what we might have expected from our cartel model (cf. above, p. 200). General Electric has been able to maintain its position because it could control entry into the light bulb industry through ownership of vital patents. The dominant anthracite coal companies have control of most of the anthracite reserves. The "Big Four" in cigarettes have prevented entry by establishing monopoly

8. Cf. Burns, op. cit., pp. 140ff.

through advertising. The depression, however, partly broke down this last monopoly by allowing the cheaper brands to establish themselves; and in recent years Philip Morris has established itself, in its turn, by a vigorous advertising campaign.[9]

The preceding paragraphs, though necessarily quite sketchy, provide a rough outline of the thesis I have tried to present in this section: that the cartel model is the most fruitful approach to economic analysis of our real world, focusing attention on the significant points of the problem. This general statement will now be amplified by applying the thesis to a specific industry and noting the particular ways in which the cartel features appear. The example I have chosen is the milk distributing industry, whose striking resemblance to our cartel model makes it truly a "textbook case."[1]

Due to the perishability and high transportation costs of their product, milkshed coöperatives, made up of thousands of members, are able to operate more or less within a closed market. The coöperative bargains collectively with the distributors and sets a price on fluid (Class I) milk. It cannot, however, control the output of its members, and therefore all milk not used in fluid form is sold as surplus (Class II) milk at prices near competitive levels. This is used for butter, cream, condensed milk, and other processed dairy products. The individual producer is either allotted a quota on which he can receive the Class I price (receiving the Class II price on everything above this quota) or he receives the Class I price on a percentage of his sales equal to the percentage of the total coöperative sales which was used for Class I purposes. Over a period of time expansion takes place as (1) new producers are attracted into the dairy industry within the existing milkshed, (2) the individual members expand their production, and (3) the higher price set by the coöperative itself extends the geographical area of the milkshed.

As the size and output of the milkshed increase, the proportion of surplus milk increases still faster. This brings down the

9. Cf. A. A. Bright and W. R. Maclaurin, "Economic Factors Influencing the Development and Introduction of the Fluorescent Lamp," Journal of Political Economy, Vol. 51 (1943) pp. 429–450; Burns, op. cit., p. 123.

1. For the following account I have drawn heavily on John M. Cassels, A Study of Fluid Milk Prices (Harvard Economic Studies No. 54) (1937), Chap. 5–6 and Appendix A; and Wm. H. Nicholls, Imperfect Competition within Agricultural Industries (1941) Chap. 10–11.

average price received by the producer. If there were completely free entry, the *average* price would tend to fall to the competitive level, with two possible results.

(1) At any time of the coöperative's existence, it is to the advantage of any individual producer to stay out of the coöperative and sell to non-participating dealers who sell primarily fluid milk. They will thus obtain a higher price than the average obtainable within the coöperative, but lower than the bargained Class I price. Note that this is also to the advantage of the non-participating distributor, since it (*a*) enables him to obtain Class I milk at a lower price than his competitor (participating) distributors and (*b*) throws a greater burden of surplus milk on them. Thus it is to the common interest of producers' and distributors' organizations to prevent free entry at both levels. When increasing amounts of surplus milk continuously lower the average price and increase the discrepancy between it and the Class I price, this unremitting pressure becomes overwhelming and, if not counteracted, causes the disintegration of the coöperative long before the "competitive" price is reached.

(2) But, of course, counter measures will be put into effect before this danger point is reached. Government assistance will be called in — health ordinances will be used to restrict the area of the milkshed and discipline recalcitrant producers. (The Rhode Island ordinance requiring milk from outside the state to be colored pink is classic.) Where this does not suffice, force can also be used (milk-dumping, for example). Action will also be directed at the non-participating distributors, for without these as an outlet the non-coöperating producer would be lost, unless he could do his own distributing. The producers' and distributors' organizations will thus try to eliminate them, either with government assistance (again via health ordinances) or pure coercion (their bottles broken or held back at the bottle-exchange, their workers beaten, and so on). Finally the Government itself is called in to fix and enforce the price, usually by arbitrating or participating in the bargaining between coöperatives and distributors and giving these results legal sanction.[2]

2. Cf. TNEC Monograph No. 32, Economic Standards of Government Price Control, Part II, "Public Pricing of Milk" for government regulation of fluid milk marketing in Oregon, California, Indiana, Wisconsin, and New York. Federal price fixing is also discussed. (pp. 57–229).

V

In conclusion it should be noted that to a large extent the general imperfect competition case described above can be dealt with by a market-sharing solution from the viewpoint of a single firm whose individual demand curve shifts to the left as new firms enter. For the following reasons, however, I believe that the cartel model is a more satisfactory analytical tool for revealing certain of the forces at work and should therefore be used in addition to the market-sharing analysis.

(1) There are some cases for which the cartel model is an exact, and not merely an approximate, description, in the sense that there is an actual body making and enforcing decisions from the viewpoint of the industry as a whole. The market-sharing analysis, with its emphasis on the individual firm, is obviously inapplicable here. In our own economy we have the examples of the milk industry, bituminous coal (Guffey Coal Act), oil (state proration laws and the Connally Act), and industries with strong trade associations. When we extend our view to the international scene, the examples become much more numerous: rubber, tin, wheat, sugar, coffee, etc. None of these can be satisfactorily analyzed from a market-sharing point of view.

(2) Even when the industry has not set up any official central office, our previous analysis has shown that through the interaction of dominant and independent firms we get approximate cartel results. Here, too, the dominant firms will make some decisions from the viewpoint of the industry. For example, they may not maximize short-run profits, in order not to attract new firms. But if we look at it from the viewpoint of each firm, the latter would not gain by foregoing its profit while the other firms retained theirs. The cartel model is necessary to bring out the nature of these industry decisions.

(3) The cartel model shows with graphic clarity the development of long-run excess capacity and overinvestment.

(4) Obviously, from a firm analysis we could not obtain the results of having new firms entering the industry and remaining idle.

(5) The cartel model reveals more precisely the inner mechanisms, forces, and conflicts leading to the disintegration of industry agreements and explains the resort to extra-economic methods of maintaining them.

UNIVERSITY OF CHICAGO   DON PATINKIN.

REGULATION OF MINIMUM RATES
IN TRANSPORTATION

SUMMARY

Importance of the problem, 206. — Congressional directives, 207. — General attitude of the Commission, 208. — Minimum rates and discrimination, 211. — Adequacy of revenue, 211. — Protection of the rate structure, 214. — Preservation of the advantages of particular modes of transport: the value standard, 220; particular agencies, 223; cost standards, 225. — Conclusions, 229.

The Interstate Commerce Commission said in 1945 that the keynote of the postwar years will prove to be a quickening and extension of competition within and among the several forms of transportation and with private carriers.[1] This prediction is already being borne out by improvements in service. Competitive adjustments have also been made in rates, and it is not unlikely that rate cutting will become severe, particularly if a substantial decline in the total volume of commercial traffic occurs. In order to check discrimination, to stabilize the credit of the transportation industries, and to curb wasteful hauling, control of minimum rates will probably be increasingly necessary. At the same time, it will be desirable to avoid an undue freezing of the rate structure, encouragement of private transportation, or nullification of the inherent advantages of the most efficient modes of transport.

Fortunately the Interstate Commerce Commission is better equipped than in the prewar decade to fix rate floors. Its authority now extends to all forms of commercial transportation except air carriers. The Commission was authorized to prescribe minimum rates for railroads by the Transportation Act of 1920, for motor carriers by the Motor Carrier Act of 1935, for water carriers by the Transportation Act of 1940, and for freight forwarders by an act of 1942. Its power to decide minima has been upheld by the Supreme Court of the United States. Not only does the Commission have broader authority, but it also has had more experience in prescribing minimum rates. However, the regulation of the lower level of rates under conditions of widespread competition between unlike agencies, when all are subject to control, is relatively new,

1. Annual Report of the Interstate Commerce Commission, 1945, p. 7.

and it can scarcely be said that consistent principles have been adopted. The Commission seems to have followed to some extent "the process of trial and error,"[2] and many of its minimum rate decisions have given rise to dissents as well as to outside criticism. The purpose of this article is to analyze the Commission's policies in the light of what appear to be rational criteria for minima.

CONGRESSIONAL DIRECTIVES

Regulation of competitive rate-making was once described by the late Commissioner J. B. Eastman as one of the most difficult problems the Commission has ever had to face. One important reason for its difficulty lies in the Congressional policies and directives bearing upon inter-carrier rate relations. The national transportation policy of 1940, which must be considered in administering every provision of the Interstate Commerce Act,[3] provides that all forms of transportation shall be impartially regulated so as to recognize and preserve the inherent advantages of each, so as to promote efficient service and sound economic conditions in transportation, and so as to encourage the establishment of reasonable non-discriminatory rates without unfair or destructive competitive practices. Section 500 of the Transportation Act of 1920 declares it to be the policy of Congress to promote transportation by water and to preserve in full vigor both rail and water transport.[4] The rule of rate making for railroads, Section 15a of the Interstate Commerce Act, directs the Commission in prescribing rates to consider the effects of rates upon the movement of traffic by the carrier or carriers for which rates are fixed, the need for adequate railway service at the lowest rates consistent with such service, and the requirement of the carriers for revenue. The rules for motor lines, water carriers, and freight forwarders — Sections 216i, 307f, and 406d — are quite similar to the railroad rule. The Long-and-Short-Haul Clause provides that the Commission shall not permit the establishment of a rate to or from a distant point that is not reasonably compensatory, prohibits recognition of merely potential water competition, and prevents the railroads

2. Annual Report of the Interstate Commerce Commission, 1939, p. 29.

3. See Eastern-Central Motor Carriers Association v. United States, 321 U. S. 194 (1944).

4. This section has not been repealed, but its special force may have been nullified by the declaration of national transportation policy in the Act of 1940.

from raising rates reduced to meet water competition without a finding by the Commission that a proposed increase rests upon changed conditions other than the elimination of water competition. The sections of the Interstate Commerce Act dealing with joint rates authorize the Commission to prescribe the terms and conditions applicable thereto, including the establishment of rail-water differentials. Differences between rail and water rates shall not be deemed to constitute unjust discrimination or an unfair or destructive competitive practice. A corresponding precept applies to variations among the rates of other types of carriers.

By the sections named above and by the provisions of the Interstate Commerce Act dealing with certificates of service and with combinations, Congress has emphasized the encouragement of competition in transportation. At the same time, it has stressed the promotion of sound financial conditions in the industry and the elimination of competitive practices inimical to good service. It has also insisted upon the establishment of reasonable rates and the preservation of the inherent advantages of each mode of transportation. These three major objectives govern minimum as well as maximum rates, and their interpretation and reconciliation constitute the Commission's basic problem in regulating minima.[5]

Since the law lays down no specific instructions for minima, the Commission has said from the beginning that "The act prescribes no different tests or standards for the determination of just and reasonable minimum rates than for just and reasonable maximum rates. And there is nothing in the act which either directs or implies that we should use tests of different general character in the determination of reasonable minimum rates than those which we have consistently and uniformly applied . . . in fixing maximum reasonable rates."[6] Criticism of this view follows later.

GENERAL ATTITUDE OF THE COMMISSION

In order to permit competition and flexibility in rate quotation, which are necessarily partly restrained when rate floors are prescribed, the Commission has generally adopted a self-denying attitude toward its authority over minima, especially as regards

---

5. Compare remarks by the late Commissioner J. B. Eastman in the Railway Age, vol. 108, pp. 717-718 (April 20, 1940).

6. Salt Cases of 1923, 92 I. C. C. 388, 410 (1924). See also New Automobiles in Interstate Commerce, 259 I. C. C. 475, 535 (1945).

railroads. In the Sugar Cases of 1922, the Commission made the following statement, which has been repeated in substance on numerous later occasions: "We believe, however, that this power should be sparingly exercised and only in cases where it clearly appears that its exercise is necessary in order that substantial public injury may be avoided."[7] This was a railroad case, but the Commission has also followed a comparatively liberal policy toward motor carriers, refusing to prescribe minimum rates except on a clear showing of need.[8] When minima have been fixed for motor carriers on a broad scale, the circumstances have been unusual.

As viewed by the Commission, reasonable minimum rates are not necessarily the equivalent of reasonable maximum rates, even though both are subject to the same directives. The common law does not hold that the lowness of a rate can constitute a wrong, and since the statutes do not otherwise provide, minimum rates may properly be lower than maximum.[9] The mere fact that proposed rates are less than reasonable maximum rates already approved does not condemn them.[1] A zone is recognized wherein the carriers, within the limits of reasonableness and prejudice, are free to reduce their rates.[2] Departure from this principle by the Commission is likely to encounter the disapproval of the Supreme Court.[3] Although regulation and competition have narrowed the zone of discretion, on high-grade commodities that load densely the spread between reasonable maxima and minima may be wide, at least in the case of railroads. Much depends upon the standard adopted for measuring the rate floor.

The Commission has suited action to words, and in thousands of instances the carriers have been permitted to reduce their charges. Most minimum rate cases come before the Commission

---

7, 81 I. C. C. 448, 472 (1923). See also 123 I. C. C. 503, 504 (1927); 195 I. C. C. 198, 202 (1933); 211 I.C.C. 307, 309 (1935); 219 I.C.C. 501, 510 (1936); 241 I. C. C. 21, 41 (1940); and 259 I. C. C. 475, 535 (1945).

8. 31 M. C C 536, 538 (1941).

9. Petroleum Products between Kansas, Oklahoma, Arkansas, Missouri, and Colorado, 245 I. C. C. 617, 634 (1941).

1. Morton Salt Company v. Alton Railroad Company, 264 I. C. C. 71, 75 (1945). It may be observed that this principle limits the practicability of the comparative method in judging the reasonableness of minimum rates.

2. Freight Forwarding Investigation, 229 I. C. C. 201, 249 (1938).

3. United States v. Chicago, Milwaukee, St. Paul and Pacific Railroad Company, 294 U. S. 499, 506 (1935).

in suspension proceedings, for which statistics have been published. It is impossible to determine from the data in the Commission's annual reports what proportion of the proposed changes in rates that represent reductions have been suspended; but the figures from 1920 to date show definitely that as the ratio of reduction cases to increase cases becomes larger, the percentage of the total number of proceedings in which the Commission denies suspension tends to rise. In 1940 the relative number of reduction cases was unusually large, and the ratio of denied to suspended and denied cases was 71 per cent.[4] During other years when proposed reductions have been especially numerous the percentage has not been so high, yet it seems reasonable to say that the Commission permits reduced rates to take effect in a very substantial number of instances.[5] If valid, this conclusion qualifies the criticism that rate regulation is too inflexible.[6]

Although the Commission appears to have exerted its authority over minimum rates with a moderate hand, it has disallowed reductions or has prescribed minima when persuasive evidence indicated that proposed or existing rates were in its opinion unlawful. A standard of lawfulness was set forth in a recent decision. "What constitutes a minimum reasonable rate is a matter to be determined in the light of the facts of record in each individual case. . . . Whether a rate is below a reasonable minimum depends on whether it yields a proper return; whether the carrier would be better off from a net-revenue standpoint with it than without it; whether it represents competition that is unduly destructive to a reasonable rate structure and the carriers; and whether it otherwise conforms to the national transportation policy and the rules of ratemaking declared in the act of 1940."[7]

4. Annual Report of the Interstate Commerce Commission, 1940, p. 131.

5. In 1938 railroads sought suspension in 137 instances of motor carrier tariffs proposing rate reductions. Of these 64 went into effect over protest and 47 were suspended. Corresponding suspensions for 46 rail tariffs protested by truckers were 15. Letter from the Chairman of the Legislative Committee of the Interstate Commerce Commission, House Committee Print, p. 5 (March 20, 1939).

6. On this point see the testimony of Commissioner J. B. Eastman in Committee on Interstate Commerce, Long and Short Haul, Senate Report No. 1768, 75th Congress, 3d Session, part 2, pp. 30–60 (1938).

7. New Automobiles in Interstate Commerce, op. cit., p. 534.

MINIMUM RATES AND DISCRIMINATION

The foregoing definition indicates that a reasonable minimum rate is not a matter of formula, and points with measured definiteness to the principal grounds upon which the Commission has based its decisions. However, discrimination is not specifically mentioned, and cases in which minimum rates have been found unlawful because of undue preference and prejudice are numerous. Reduced rates proposed by any type of carrier will be condemned if they unduly favor certain shippers (or classes of traffic) as compared with others similarly situated.[8] Rail and forwarder minima will be disallowed if they discriminate against interstate commerce.[9] Here the primary consideration is whether intrastate rates unduly prefer intrastate as compared with interstate shippers, and the question whether the rates are reasonable per se is generally irrelevant. Motor and water minima are not subject to this rule, for the Interstate Commerce Act definitely provides that the Commission may not prescribe motor and water rates in order to remove discrimination against interstate commerce.

ADEQUACY OF REVENUE

The first consideration mentioned in the standard of lawfulness referred to above is whether a rate yields enough revenue. This factor frequently applies when the rate also represents destructive competition, but the revenue test governs irrespective of competitive repercussions.[1] As a rule, the adequacy of a rate is determined with reference to the cost of the service, yet in some instances proposed reductions have been condemned on the ground that the decline in unit revenue from the low rates seems likely to more than offset the gain from additional traffic.[2] However, cases of this kind involve the substitution of the Commission's judgment for that of management, and the Commission seems inclined to resolve the doubt in favor of the companies.[3]

8. Whisky from Frankfort, Ky., to Lawrenceburg, Ind., 241 I. C. C. 465, 469 (1940).

9. Southwark Manufacturing Company v. Pennsylvania-Reading Seashore Lines, 231 I. C. C. 755, 767 (1939).

1. Ex-Lake Grain to North Atlantic Ports, 235 I. C. C. 415, 428-429 (1939).

2. All Freight between Portland and Seattle, 238 I. C. C. 729, 734 (1940).

3. Lumber from Pacific Coast to Eastern Points, 210 I. C. C. 317, 346 (1935). See also 227 I. C. C. 485, 493 (1938).

Whether a rate is reasonably compensatory is an important, if not the predominant, element to be considered in determining lawful minima.[4] A reasonably compensatory rate has been defined as "one which is remunerative, i.e., covers the out-of-pocket costs . . . of handling the particular traffic under consideration, including a proper return on investment."[5] Out-of-pocket costs, as referred to in this definition, embrace those expenses which are added over a period of time by an increase, or reduced by a decrease in traffic, can be assigned to specific traffic movements, and are directly affected by the use of carrier facilities.[6] This seems to mean the economist's marginal cost, but it appears that the Commission has usually identified out-of-pocket cost with average variable cost.[7] A return on the investment in road and equipment required in the long run to accommodate added traffic is omitted from out-of-pocket costs, although the Commission pointed out that from 50 to 70 per cent of the value of the property used in freight transportation is affected by traffic volume.[8]

Minimum rates based upon out-of-pocket costs in the short run are not reasonably compensatory, and will not be prescribed. With reference to rates that are remunerative only in the narrow sense, the Commission has said, "Defendants also invoke the familiar argument that shipments transported under the all-freight rates are added traffic, that is, freight which, except for the maintenance of the challenged rates, would not move by railroad at all, and that accordingly, if the resulting revenue is sufficient to cover out-of-pocket costs plus at least a small contribution to the overhead or system costs, the rates are not too low, but on the contrary, they actually lighten the burden of other traffic. Whatever validity the added-traffic theory of rate-making may have had some years ago, we think it has very little today. . . . The addition of another car to a freight train doubtless makes but a slight increase in the cost of operating the train, so slight indeed that the out-of-pocket cost of moving any particular car may be calculated, on

4. Freight, All Kinds, Lincoln, Omaha, and Nebraska Points, 32 M. C. C. 339, 344 (1942).

5. New Automobiles in Interstate Commerce, op. cit., p. 538. See also Petroleum from South Atlantic Ports to Southeast, 245 I. C. C. 23, 29 (1941).

6. New Automobiles in Interstate Commerce, op. cit., p. 500.

7. See Wilson, G. L., and Rose J. R., "Out-of-pocket Cost in Railroad Freight Rates," this JOURNAL, vol. 60, pp. 546–560 (August, 1946).

8. New Automobiles in Interstate Commerce, op. cit., p. 528.

that basis, as practically zero. However, overhead costs are no less real than direct transportation costs, and if all of the cars in the train, or any considerable number of them, are moving at submaximum rates, as must often be the case today, then clearly there is no room for application of the added-traffic theory."[9] This was a railroad case, but the same principle has been applied with even more certainty to highway transportation. The Commission has pointed out that the dangers of the out-of-pocket basis are, if anything, greater for motor carriers than for railroads.[1] The former have less noncompetitive traffic than the latter.

Not only has the Commission rejected strictly out-of-pocket costs as a measure of prescribed minima, it has frequently included in the term "reasonably compensatory" fully distributed or average costs, i.e. operating expenses plus a normal return on investment.[2] In fact, as indicated in the next section, proposed reductions that cover out-of-pocket costs plus a proportionate share of the overhead have sometimes been condemned.

Although the words "reasonably compensatory" and "proper return on investment" have been interpreted as referring only to cost,[3] it is possible that the Commission thinks they depend in part upon the value of the service and may call for a more than proportionate contribution to the overhead. The Commission has declared over and over that the tests of reasonable minimum rates are the same as those governing maxima, which give considerable weight to value factors. It has spoken of rates that are "reasonably compensatory from the standpoint of cost of service,"[4] and in one case said, "In determining what are reasonable minimum rates, we may take into consideration, to the extent that conditions will permit, the value as well as the cost of the service."[5] In other words, what constitutes a reasonably compensatory rate may

9. Middle Atlantic States Motor Carrier Conference v. Central Railroad Company of New Jersey, 232 I. C. C. 381, 391 (1939). See also All Freight from Eastern Ports to the South, 251 I. C. C. 361, 367 (1942).

1. Refrigerator Material from Memphis, Tenn., to Dayton, Ohio, 4 M. C. C. 187, 189 (1938).

2. Petroleum from South Atlantic Ports to Southeast, op. cit.

3. Witters, Myron, "A Study of Minimum Reasonable Rates," I. C. C. Practitioners' Journal, vol. 13, pp. 438-509, 441 (March, 1946).

4. Petroleum Products between Kansas, Oklahoma, Arkansas, Missouri, and Colorado, 245 I. C. C. 617, 634 (1941).

5. Candy from Reading, Pa., to Baltimore, Md., 237 I. C. C. 89, 95 (1940). See also Cotton Piece Goods in the South, 234 I. C. C. 525, 539 (1939); and Petroleum from South Atlantic Ports to Southeast, op. cit., p. 24.

depend upon what the traffic will bear in particular cases. Sometimes prescribed rates may equal or even exceed average costs. At other times they lie closer to out-of-pocket costs.

Whether the term "reasonably compensatory" refers only to cost is, of course, relatively unimportant. The significant fact is that the Commission has employed the value of the service as one basis for minimum reasonable rates. As indicated, in some instances minima have been prescribed above full costs. In practically all cases they have been fixed at a level higher than necessary from a strictly compensatory point of view.

Rejection of the out-of-pocket cost principle for minima is consistent with the dictum of the Supreme Court in the Northern Pacific Case[6] and with the policy followed by the Commission in administering Section 4 of the Interstate Commerce Act;[7] but from the point of view of a particular carrier, it is sound only on the assumption that lower rates will not reduce the average cost of service. When plant has already been constructed, insistence upon full costs may be improper. For this reason the Commission often allows the carriers to quote rates that approach out-of-pocket costs in the narrow sense, if reductions are necessary to meet compelling competition.[8] However, rates must not cut below out-of-pocket costs, and must be no lower than necessary to meet competition.[9]

PROTECTION OF THE RATE STRUCTURE

The key to the Commission's prescription of rates that are more than strictly remunerative lies in the need to protect the rate structure. Here the question is not whether minimum rates are profitable to a particular carrier, but whether they threaten to

6. Northern Pacific Railway Co. v. State of North Dakota, 236 U. S. 585, 597 (1915).

7. Transcontinental Cases of 1922, 74 I. C. C. 48, 71 (1922). It may be noted that the term "reasonably compensatory," as used in Section 4, involves much more than the netting of some profit, and that it does not necessarily provide the test for reasonable minima. Ex-Lake Grain to North Atlantic Ports, op. cit., p. 429.

8. All Freight from Eastern Ports to the South, op. cit. See also Naval Stores from Mississippi to Gulf Ports, 235 I. C. C. 723 (1940).

9. Molasses from New Orleans, La. to Peoria and Pekin, Ill., 235 I. C. C. 485, 502 (1939); Mineral Water from Hot Springs to North Atlantic Ports, 245 I. C. C. 535, 541 (1941).

deplete revenues and endanger service on a broad basis.[1] The interest of the individual carrier must sometimes give way to that of the group, even though the rates proposed by that carrier are justified from its point of view. If one carrier is permitted to charge rates that are profitable only in the strictest sense, other carriers are entitled, in the opinion of the Commission, to the same privilege. All are prone to reduce rates to the lowest levels possible without direct loss of revenue.[2] And if the competitive traffic amounts to a large percentage of total business, as is particularly true of motor carriers, the result in the long run is likely to be inadequate earnings and impaired service in general. Under such circumstances the Commission will insist upon rates substantially in excess of out-of-pocket costs in the short run.

With reference to rates that endanger general credit, the Commission has said, "Some persons contend that any rate that produces earnings in excess of the out-of-pocket costs of handling traffic without allocating thereto anything for other costs or return on investment is a minimum reasonable rate. Out-of-pocket costs may be indicative of the extent to which carriers may go in meeting competition without financial loss, and we have found such rates to be compensatory and not unreasonable per se when made to meet compelling competition. It is apparent, however, that if all or a large proportion of railroad rates were brought down to such a level, the vitality of the railroad system would be destroyed. The Transportation Act of 1940 imposes on us the duty to scrutinize rates purportedly made on the out-of-pocket cost or minimum-rate theory to meet alleged competition and to reject them even if they yield something above all costs when they are less than prevailing reasonable rates prescribed by us, in the absence of proof that they are required to meet the compelling competition, and that they will not foster unsound economic conditions in transportation among the several carriers or otherwise contravene the policy declared by the Congress in the 'national transportation

1. A few cases involving the rate structure have been decided for the purpose of maintaining structural relationships, rather than of conserving revenue. Cotton, Wool, and Knitting Factory Products, 220 I. C. C. 189, 191 (1937).

2. An excellent illustration of progressive rate-cutting occurred between railroads and truck-water routes in transporting citrus fruit from Florida to North Atlantic ports. Citrus Fruit from Florida to Baltimore, Md., 237 I. C. C. 245 (1940).

policy.' In the absence of convincing evidence that special rates are justified by the facts and circumstances in particular situations, every rate should bear its fair share of the transportation burden."[3]

The power of the Commission to protect the rate structure and prevent the gravitation of rates to the lowest possible levels has been upheld by the courts. In construing the Act of 1920, the Supreme Court said that "the interests of the individual carrier must yield in many respects to the public need,"[4] and in the Anchor Coal case a district court declared that the Commission "has the right to prescribe minimum rates . . . to prevent ruinous rate wars and to guarantee reasonable earnings, not only to the carrier affected but also to competing carriers, who may labor under a higher cost of doing business."[5] In a later case the Supreme Court indicated by dictum that the Commission cannot prescribe minima unless rates are so low as to cast a burden on other traffic;[6] and in the Milwaukee case the Court pointed out that every rate change disrupts rate structures to a certain extent, that the fear of reduced earnings through rate wars engendered by the publication of low rates may prove to be illusory.[7] However, the order of the Commission in the Milwaukee case was inadequately buttressed by evidence, and soon afterward the Court upheld a similar, well-supported order.[8] In the Scandrett case the Supreme Court confirmed a decision of a lower court that sustained the Commission in condemning proposed rates that were compensatory, considering all costs.[9] What was said in the Anchor Coal case has come to be settled law.[1]

Minimum rates to protect the rate structure have generally been prescribed for railroads or motor carriers. Few water carrier cases have yet arisen, although in 1940 the United States Maritime Commission fixed minima to prevent rate wars in the intercoastal

3. All Freight from Eastern Ports to the South, op. cit.

4. United States v. Illinois Central Railroad Company, 263 U. S. 515, 525 (1924).

5. 25 F. (2d) 462, 471-472 (1928).

6. Texas and Pacific Railway Co. v. United States, 289 U. S. 627, 633 (1933).

7. United States v. Chicago, Milwaukee, St. Paul and Pacific Railroad Company, 294 U. S. 499 (1935).

8. Youngstown Sheet and Tube Company v. United States, 295 U. S. 476, 480 (1935).

9. 32 F. Supp. 995 (1940); 312 U. S. 661 (1941).

1. New Automobiles in Interstate Commerce, op. cit., p. 535.

trade.[2] After the Transportation Act of 1940 became effective, these minima were modified by the Interstate Commerce Commission in order to permit the intercoastal carriers to meet competition from all-rail routes.[3]

One of the first instances involving minimum rates for railroads occurred when the carriers serving competitive producers in Kansas, Louisiana, Michigan, and New York were quoting unduly low rates in affording outlets for salt.[4] In this case the rate was condemned as too low, not on the ground that it would cause loss to the publishing carrier, but because of its harmful effects upon the rate structure in general. Soon afterward, reduced rates on iron ore were disapproved as a threat to railroad earnings.[5] During the depression of the 'thirties proposed rates on gasoline and coal were disallowed for similar reasons.[6] In these cases competition from water carriers was a factor. More recently, low rates on export grain were set aside as dangerous to the rate structure.[7] It may be observed that these and other examples which could be cited provided for minimum rates only on specified commodities.

Regulation of the minimum rates of motor carriers has been especially important because of the highly competitive character of the industry, large numbers of operators, different service obligations, varying costs, presence of three types of carriers, lack of experience in fixing rates, disunity of action among the companies, and other factors. These conditions have led to irresponsible rate cutting and wide variations in rates. In order to stabilize rates and to create greater uniformity, common-carrier trucking associations have published minima, commonly called "stops," but these minima have led to difficulties. Since they have been based upon railroad rates and classifications, the effect has been to deprive commodities rated lower than the minimum classes of their more favorable transportation characteristics.[8] Another disadvantage

2. 2 U.S.M.C. 285 (1940).

3. Intercoastal Rate Structure, 253 I. C. C. 331 (1942).

4. Salt Cases of 1923, op. cit.

5. Ex-Lake Iron Ore from Chicago to Granite City, 123 I. C. C. 503, 505 (1927).

6. Gasoline from San Francisco Bay Points to Ogden, Utah, 198 I. C. C. 683, 695 (1934); Coal from Kansas and Missouri to Missouri via Missouri-Kansas-Texas Railroad, 198 I. C. C. 535, 538 (1934).

7. Export Grain from Central Freight Association Territory to North Atlantic Ports, 235 I. C. C. 655, 672 (1940).

8. Minimum Class Rate Restrictions, Central and Eastern States, 44 M. C. C. 367 (1945).

is that the minima frequently do not apply to all lines, some carriers maintaining stops higher than others. This often causes shippers to pay through rates based upon the high stops. The shipper might elect another route over which a lower minimum governs, but it would be necessary to examine many tariffs in order to determine the most economical route. All carriers might be forced to charge the same minima, as has sometimes been done, but this would tend to ignore substantial variations in carrier costs.

Minima have been prescribed for motor carriers on a territorial basis for a large number of companies, as well as on a more restricted basis. Territorial decisions have fixed rates on commodities in general, with exceptions, in the Middle Atlantic States, New England, Central Territory, Trunk Line Territory, and Midwestern Territory.[9] The rates put into effect were of an emergency character, being approved by the Commission in most instances as proposed by the associations of carriers, without careful analysis of costs or particular services. Many of the rates were the same as, or definitely related to, railroad rates and tended to ignore motor carrier costs. In so far as influenced by such costs, the rates were generally based upon the average costs of all carriers.

These summary decisions of the Commission were criticized on the ground that shippers were unfairly deprived of low rates and that the discretion of management was unduly restricted, but the Commission justified its action by the situation facing the motor-carrier industry, which was found to be too critical to permit the delays involved in a more careful determination. While the carriers had been slashing rates, costs had been rising, so that revenues were generally barely equal to or less than operating expenses. As a result, adequate and dependable motor-carrier service was endangered.[1]

When passing judgment on the Commission's procedure, the impartial observer, though conscious of the strong efforts of the organized carriers to protect themselves even from legitimate

9. Middle Atlantic States Motor Carrier Rates, 4 M. C. C. 68 (1937); New England Motor Carrier Rates, 8 M. C. C. 287 (1938); Central Territory Motor Carrier Rates, 8 M. C. C. 233 (1938); Trunk Line Territory Motor Carrier Rates, 24 M. C. C. 501 (1940); Midwestern Motor Carrier Rates, 27 M. C. C. 297 (1941). See also California Motor Carrier Rates, 41 M. C. C. 19, 41 (1942).

1. See for example 4 M. C. C. 68, 77 (1937); and 8 M. C. C. 233, 253 (1938).

competition, should take into consideration the fact that the minimum rates were looked upon as standards from which subsequent departures could take place.[2] Motor operators were expected to ascertain costs and to develop a system of rates related as closely as possible thereto. Individual carriers able to justify rates lower than the territorial minima were to publish such rates, although wholesale departures from the general standards would obviously defeat their purpose.[3] Those proposing lower rates were merely required to show their rates to be reasonably compensatory.

On many occasions the Commission has modified the territorial decisions, and in 1943 it vacated or suspended its orders in the proceedings for Trunk Line, Central, New England, and Midwestern territories, having found that the rate cutting in practice at the time the cases began had practically ceased.[4] Moreover, in 1945 class rate restrictions set up by the carriers were reduced.[5] The Commission has declared that when proposed individual motor carrier rates are reasonably compensatory, only the most impelling reasons justify the maintenance of minima.[6]

In the more restricted proceedings dealing with minimum rates for motor carriers, decision generally turns on considerations such as whether the rates are above out-of-pocket costs,[7] whether the rates are necessary to meet competition,[8] and how the rates compare with the charges of other motor carriers.[9] Sometimes, however, rates are condemned as threats to the rate structure and service in general.[1]

2. 27 M. C. C. 297, 320 (1941).

3. McCormick Transfer, Commodities in Massachusetts and New Hampshire, 22 M. C. C. 385, 388 (1940).

4. Railway Age, vol. 115, p. 705 (October 30, 1943).

5. Minimum Class Rate Restrictions, Central and Eastern States, op. cit.

6. Freight, All Kinds, Lincoln, Omaha, and Nebraska Points, 32 M. C. C. 339, 344 (1942); Minnesota-North Dakota Motor Carrier Rates, 43 M. C. C. 289, 304 (1944).

7. 4 M. C. C. 187, 190 (1938).

8. 4 M. C. C. 641, 649 (1938).

9. 24 M. C. C. 419, 422 (1940).

1. Commodities from Kansas to Illinois, Missouri, and Oklahoma, 44 M. C. C. 90, 93 (1944). See also Morton Salt Company v. Alton Railroad Company, 264 I. C. C. 71, 88-91 (1945).

Preservation of the Advantages of Particular
Modes of Transport

Another objective of minimum-rate regulation besides the protection of the rate structure is the promotion of an economical division of traffic among the transport agencies, as required by the national transportation policy.[2] The two ends are not the same. Within the limits of what the traffic will bear, including potential competition from private means, it is possible to conserve the credit of all carriers and yet permit relationships among the rates of particular agencies which will not direct traffic according to the inherent advantages of different carriers. The desideratum, of course, is to accomplish both purposes at the same time in such a manner as to encourage healthy competition.

Preservation of the inherent advantages of each mode of transportation as an objective of minimum-rate regulation was expressed by Commissioner Eastman in the following words: "There are, I believe, sound grounds for holding that we were given the minimum-rate power for the purpose of ... promoting within reason the use, to the extent that our jurisdiction permits, of the different modes of transportation for the services to which they are economically best fitted and discouraging their use under reverse conditions. ..."[3] In the Scandrett case the Court stated that it is appropriate for the Commission in fixing rates to take into account the apportionment of traffic among competing agencies.[4]

How can minimum rates be fixed so as to preserve inherent advantages and at the same time conserve credit? According to one view, minima should be based upon the value of the service, not necessarily reflecting costs at all.[5] Under this plan a relationship is fixed between the rates of competing agencies which will neable all to participate in the traffic, taking into account costs and service advantages. After this relationship has been established, the rate

---

2. This question is discussed in Locklin, D. P., "Transport Coordination and Rate Policy," Harvard Business Review, vol. 15, pp. 417–428 (Summer, 1937). See also Oppenheim, S. C., The National Transportation Policy and Intercarrier Competitive Rates (1945).

3. Naval Stores from Mississippi to Gulf Ports, op. cit., pp. 737–738 (1940).

4. 32 F. Supp. 995 (1940).

5. See Annual Report of the Interstate Commerce Commission, 1939, p. 28; and Naval Stores from Mississippi to Gulf Ports, op. cit., p. 738 (1940).

level is set at least as high, if not higher than the costs of the most inefficient agency, provided the traffic permits. To illustrate, in the Naval Stores case the Commission approved a rail-truck differential in favor of the railroads of about 2 cents per 100 pounds on rosin from Hattiesburg to Gulfport, Miss., and also permitted a proposed rail rate of 6 cents, which was approximately equal to railroad out-of-pocket costs plus a normal return on investment.[6] By reason of service advantages, it was found that trucks could successfully compete on the basis of a spread of 2 cents, and that a truck rate of about 8 cents would more than cover all motor carrier costs.

The principal argument in favor of the value standard is that it promotes the financial stability of all modes of transportation, provided there is enough business for all. It assumes that competition forces rates down toward an out-of-pocket basis on a very substantial portion of traffic, and that recovery of full costs requires rates on other traffic in excess of variable expenses. Excepting pipe lines and certain types of water carriers, no means of transportation has a cost advantage on its entire business.

Although this argument carries weight, it overlooks what appear to be more forceful counter considerations. The value basis implies, or at least does not deny, the desirability of minima on the better-paying traffic at a level suited to the needs of the high-cost agency. Such minima are ordinarily unfair to shippers, who are entitled to service at a price that would be forthcoming under competition. Rates should be no higher than the most efficient carrier could afford to charge, were the traffic involved its only business; and as a rule they should be even lower, since the carrier will profit from other business.

A related objection to the value basis is that the sharing of traffic among the low- and the high-cost agencies tends to increase the costs of the former, unless it enjoys enough traffic to operate at the point of greatest efficiency. Insofar as costs rise, inherent advantages are not preserved and traffic is allocated in an arbitrary manner. There exists no objective guide for determining how much business each carrier should receive.

Another disadvantage of minima adjusted according to the value of the service is that they invite contract and private transportation. This was a factor taken into account by the Com-

6. Ibid., p. 735.

mission when permitting the railroads to reduce rates in the naval stores case.[7] Encouragement of contract and private service makes more difficult profitable operation by the common carrier. Within limits, the latter should be encouraged, as provided for in the Interstate Commerce Act and as recognized in various decisions by the Commission.[8] Section 218(b) of the Act specifies that the minimum rates of contract carriers "shall give no advantage or preference to any such carrier in competition with any common carrier by motor vehicle." The contract carrier, of course, performs a legitimate function, and the rates it charges may not be condemned merely because they are lower than common-carrier rates.[9]

Notwithstanding the objections to the value-of-the-service basis, the Commission seems to have adopted it in certain cases. One of the most criticized decisions involved minimum rates on petroleum products moving by rail, water, and highway between Washington, Oregon, Idaho, and Montana.[1] Here rates advocated by the railroads were found to be compensatory, considering all costs, but inasmuch as the proposed rates were lower than the rates being charged by the other agencies, the Commission prescribed rail minima near the cost of handling the traffic by river and truck.[2] It stated that it would approve minima which "will promote a somewhat healthier degree of prosperity for all the carriers concerned, by rail, by highway, and by water."[3] Since the railroad was the low-cost carrier, this finding "held an umbrella" over the high-cost carriers and forced shippers to pay excessive rates. It has been suggested that the cost data in this case were none too satisfactory, that the costs referred to did not include taxes and a return on investment, and that the railroad may have been in reality the high-cost agency.[4] However, the Commission failed to say that the proposed rates would *not* yield a reasonable return.

7. Ibid.

8. Contracts of Contract Carriers, 1 M. C. C. 628, 629 (1937).

9. New England Motor Rate Bureau v. James A. Lewers and John H. McCauley, 30 M. C. C. 651, 664 (1941). See also Contract Minimum Charges from and to Baltimore, Md., 32 M. C. C. 273, 283 (1942).

1. Petroleum between Washington, Oregon, Idaho, and Montana, 234 I. C. C. 609 (1939).

2. The key-point rate advocated by the railroads was 25 cents, but the Commission prescribed a rate of 28.5 cents.

3. Ibid., p. 637.

4. Locklin, D. P., "Rates and Rate Structure," in National Resources Planning Board, Transportation and National Policy, pp. 87-128, 119 (1942).

In fact, it disapproved rates from northern Montana points even though it stated that the proposed rates would afford a substantial profit.[5]

Another decision of similar nature dealt with petroleum traffic in the Southeast.[6] Here the rates proposed by the railroads, the low-cost agency, would not cover operating expenses plus a full return on investment, but they were definitely compensatory and were disallowed primarily because of being substantially below truck costs. As in the first petroleum case, minima were prescribed for the purpose of enabling all agencies to share in the traffic. In a dissenting opinion Commissioner Splawn pointed out that the railroads were deprived of traffic that would enable them to profit from the fuller utilization of plant.

The Commission has also followed the value standard when permitting motor carriers to adopt railroad classifications.[7] Irrespective of the fact that there appears to be relatively little room for the use of demand factors in the adjustment of motor rates, the Commission said in the Central Territory case, "It requires no great penetration, however, to perceive that such a method of making motor-carrier rates cannot produce satisfactory results so far as earnings are concerned, unless the rates of competitors are constructed in a similar manner. . . . It may be that it will ultimately be found that both railroad and motor-carrier rates should conform to cost-of-service principles, but it is clear that as matters now stand motor-carrier rates cannot reasonably be constructed without regard to the competitive rates of other carriers."[8] Had the Commission been prescribing motor-carrier rates under more normal conditions, it might have been less willing to approve the railroad scheme.

Instead of being so adjusted that all agencies, including high-cost ones, can share in the traffic, minimum rates may be fixed with reference to the agency for which the rates are quoted. This does not necessarily mean that only one agency can participate, but it does mean that rates are determined primarily from the standpoint of their lawfulness in relation to a particular agency. Although

5. Petroleum between Washington, Oregon, Idaho, and Montana, op. cit., p. 639.

6. Petroleum from South Atlantic Ports to Southeast, op. cit.

7. Rates between Arizona, California, New Mexico, and Texas, 3 M. C. C. 505, 533 (1937).

8. 8 M. C. C. 233, 249-250 (1938).

occasional deviations have occurred, most of the Commission's decisions seem to conform to this principle, at least since 1940. The following excerpt from the Seatrain Lines case presents the Commission's view: "In connection with these contentions it is of interest that the rule of rate making in section 15a has been recently amended so as to require us, in the exercise of our authority to prescribe just and reasonable rates, to 'give due consideration, among other factors, to the effect of rates upon the movement of traffic *by the carrier or carriers for which the rates are prescribed. . . .'* This admonition is repeated in sections 216(i) and 307(f), containing rules for rate making relating to motor and water carriers, respectively. The words which have been italicized for emphasis are of particular interest. They were not in either of the bills originally passed by the Senate and the House of Representatives but were added by the committee of conference. Their meaning, supported also by the legislative history, seems to be that no carrier should be required to maintain rates which would be unreasonable, judged by other standards, for the purpose of protecting the traffic of a competitor."[9] In this instance, proposed increases in the rates of Seatrain were denied on the ground that they were not justified by higher costs and would deprive shippers of a superior service.

Although the foregoing quotation is inconsistent with the decisions in the petroleum and certain other cases, in 1945 in New Automobiles in Interstate Commerce the Commission explained its minimum rate policy at some length, and declared the view set forth in the quotation to represent its present position.[1] The particulars of the automobile case are of interest. It embraced all rates, charges, rules, regulations, and practices in the United States subject to parts I and II of the Interstate Commerce Act applying to the interstate transportation of new automobiles by railroad in carloads, by truck-away, tow-bar, drive-away, or by any of these modes in conjunction with water carriage.

The National Automobile Transporters Association, the members of which were adversely affected through rates maintained by the railroads, requested the Commission to prescribe minimum rates higher than existing charges and on or close to reasonable maximum rates for both railroads and motor carriers. Existing

9. Seatrain Lines v. Akron, Canton and Youngstown Railway Company 243 I. C. C. 199, 214 (1940).

1. 259 I. C. C. 475, 537 (1945).

rail rates varied according to competitive conditions from 30 to 70 per cent of first class rates, and the Association proposed rates beginning at 50 per cent of first class for the shortest distances and gradually rising to first class for distances exceeding 500 miles. It also urged the Commission to prohibit any reductions below present rates to meet carrier competition, except upon a special showing. It asked that contract carrier rates be increased on the ground that they depress rail rates and thus divert traffic from common carriers by motor vehicle.

The railroads opposed the prescription of higher minima. They argued that their present rates are compensatory, that it is necessary to publish rates substantially below truck rates in order to handle any considerable portion of the business, and that they should be left free to quote such rates as are necessary to obtain their share of the traffic. Chrysler Motors also objected to the prescription of minimum rates on the ground that competition is beneficial, since it has forced a reduction in rates throughout the country. The Nicholson Universal Steamship Company argued that if a relation between the rates of the various forms of transportation is fixed, a differential should be established in favor of joint water-rail and joint water-truck rates to reflect the inferior service and low costs of joint transportation.

The Commission refused to prescribe the proposed minima. According to its cost studies, the rates of the several forms of transportation were compensatory, except in scattered instances, and the financial stability of the carriers was not threatened. It stated that within reasonable limits the public was entitled to rate reductions brought about by competition. The Commission also said that the rates advocated by the Association would spread an umbrella over motor carriers for movements exceeding 200 miles, because of the door-to-door service and other advantages of such carriers. Long-distance traffic would be diverted from the railroads, despite the circumstance that from the standpoint of fully distributed costs the rail carriers were the fittest agency beyond 200 miles. The Commission found no warrant for believing that rail rates should be held up in order to preserve a motor-rate structure, or vice versa.

The automobile case seems definitely to set aside the umbrella principle of rate making. In fact, decisions conforming to the Seatrain doctrine were handed down prior to the Transportation

Act of 1940. In a third petroleum case the Commission said, "If the costs of one transportation agency are so high as to prevent profitable operation at rates which permit the competing agency to perform satisfactory service to the public and to earn a good profit, it seems obvious that the high-cost agency in meeting the rates of the low-cost agency is attempting to compete on a nonprofit basis. To direct the low-cost agency in these circumstances to increase its rates would be to disregard the admonition of both the Interstate Commerce Act and the Motor Carrier Act to give due consideration 'to the need in the public interest of adequate and efficient ... transportation service at the lowest possible cost consistent with the furnishing of such service.' It would be regulation in the interest of the high-cost agency rather than in the public interest."[2]

If minimum rates are fixed with reference to the agency for which the rates are quoted, they should be based principally upon cost, with perhaps some allowance for value considerations on the part of the low-cost carrier. However, more than one cost level is said to be defensible.[3] According to one theory, each carrier should be permitted to reduce rates to that level necessary to attract traffic, provided the rates more than cover out-of-pocket expenses. The principal argument in behalf of this basis is that it allows a relatively large degree of flexibility and competition in rate making. Contrary arguments are that minimum rates extensively determined with reference to out-of-pocket costs undermine carrier earnings and distribute traffic, not according to relative economy and fitness, but according to the agency having the largest backlog of non-competitive business. Traffic moves by the carrier having the lowest out-of-pocket costs in the short run, and that carrier is not necessarily the most economical in the long run. For example, in the case of New Automobiles in Interstate Commerce it was found that the truck-away carrier was the most efficient agency for transporting automobiles up to about 160 miles, considering all costs, but that in most instances the railroads had the lowest out-of-pocket expenses.[4] Many similar cases could be cited.

2. Petroleum and Petroleum Products from California to Arizona, 241 I. C. C. 21, 43 (1940). See also Naval Stores from Mississippi to Gulf Ports, op. cit.

3. See Eastman, J. B., "The Adjustment of Rates between Competing Forms of Transportation," American Economic Review, vol. 30, part 2, pp. 124-129 (March, 1940).

4. 259 I. C. C. 475, 567 (1945).

A second alternative in applying the cost principle is to base minima upon the fully-allocated costs of each agency. On traffic handled at about the same cost, rates for the different types of carriers are equal and traffic is distributed largely according to service advantages. On other traffic, minima are unequal; and if services are comparable, the traffic moves by the agency of lowest cost. In contrast with the out-of-pocket plan, this procedure assures the survival of the most efficient agencies, if plant is properly adjusted to the demand for service. Another advantage is the allocation of traffic according to a reasonably definite standard. So long as an agency is allowed to depart from average cost in publishing rates, judgment is involved in deciding the extent of departure and the share of traffic to be moved by that agency.[5] Discretion is of course required in the ascertainment of costs, but there is less need for the exercise of opinion than when adjusting rates on the basis of demand.

Still another advantage of requiring each agency to charge rates measured by its own costs is the reduction of wasteful transportation. Permitting a high-cost carrier to divert traffic from a low-cost one often results in a loss equal to the difference between the direct costs of the two, as when both have excess capacity. Only if the additional expense of handling traffic by the high-cost agency is less than the corresponding expense of the low-cost agency is it economical for traffic to move via the first. This situation might occur when the low-cost agency is fully utilized. But if both are operating at capacity, rates should be based upon average costs.

Although the average-cost plan has substantial advantages, it permits competition only insofar as fully-allocated costs are equal, and it tends to cause a high degree of carrier mortality.[6] If full costs were strictly applied to railroads, numerous rates under Section 4 would be precluded, many charges on less-carload traffic, and possibly most passenger fares. Since the policy of Congress has been to encourage competition and to protect competing agencies, at least to some extent, general approval of this plan is doubtful. The decision of the Supreme Court in the Eastern-

5. See remarks by Commissioner Eastman in Naval Stores from Mississippi to Gulf Ports, op. cit., pp. 734-739.

6. Annual Report of the Interstate Commerce Commission, 1939, p. 28.

Central case apparently condemns the determination of minimum rates altogether on the basis of costs and without consideration of competitive factors.[7] In this case certain truckers, in order more effectively to compete with railroads, proposed rates based upon a volume minimum, rather than a truckload minimum weight. These rates were disapproved by the Commission on the ground that the motor carriers failed to show that they could handle a volume minimum weight at less cost per 100 pounds than a truck-load minimum. The Court objected to this finding, saying that the Commission may not decide the question solely on the basis of a cost rule formerly applied to railroads. Although the Court stated that more adequate evidence might support the Commission, it insisted upon the necessity for considering the competitive relationships between railroads and motor carriers. Three justices dissented, contending that the Commission did not adopt an inflexible cost principle, and that pursuit of the policy of the majority decision would force the Commission to embark upon far-reaching investigations when questions of rail-motor rate relationships arise.

A third cost plan — a variation of the second — is to base minimum rates upon the fully-allocated costs of the low-cost agency and to grant other carriers limited opportunity to meet the minima. This policy differs from the value standard in that rates are never higher than the average costs of the most efficient means. It differs from the measurement of minima by out-of-pocket costs in that rates are kept high enough to enable at least one agency to operate at a profit, on the assumption that the agency's average costs are below the out-of-pocket costs of competitors on a sufficient number of commodities and hauls. It differs from the policy of basing the rates for each mode of transportation upon its own average costs in that the traffic priced by the low-cost agency is shared with competitors so long as no rate falls below a remunerative level.

Regulation of minimum rates according to this third cost plan means the movement of traffic by some carriers at less than its full cost. Unless safeguarded, this is a dangerous policy and might make necessary the upward adjustment of certain rates in order to enable any one agency to earn a reasonable return. Nevertheless,

7. Eastern-Central Motor Carriers Association v. United States, 321 U. S. 194 (1944).

the plan has several advantages. First, it conforms to a large extent with the rate-making procedures of the past. Many cases can be cited in which the Commission has permitted a high-cost carrier to meet lower rates, especially in connection with the Long-and-Short-Haul Clause where roundabout lines compete with direct lines. Second, it encourages competition, thus conforming to a definitely expressed Congressional policy. Minima based upon the full costs of each agency allow competition only insofar as average costs are equal. Third, under some conditions it makes possible more economical transportation. As explained above, where a high-cost agency's plant is not fully utilized, it may be desirable to permit that agency to quote rates that approach direct costs. This is frequently the situation in the case of railroads. Fourth, it may afford a high-cost and already established means of transportation a better opportunity to remain in business. Under some conditions the diversion of traffic from a low-cost carrier to a less economical one may not seriously harm shippers served by the former, and at the same time may make it possible for the latter to operate.

The right of a high-cost agency to meet the rates of a low-cost carrier should be limited. In no case should the former be permitted to reduce its rates below strictly out-of-pocket costs, or to engage in unduly wasteful transportation. The high-cost agency should make its rates no lower than necessary to meet competition. Where it possesses service advantages, its charges should be above the rates of the controlling carrier.

In order to facilitate reduced costs for the low-cost agency, and thus the lowest rates for the public, it may be desirable to permit that agency, within limits, to quote promotional rates. By so doing, average costs can sometimes be reduced through the utilization of such excess capacity as may exist and through the promotion of such economies of large-scale production as may be realizable. But rates should not be reduced merely to divert traffic from another carrier, and plant should not be expanded to handle additional traffic unless the added business will cover average costs.   .

CONCLUSIONS

A rigid minimum-rate formula is probably undesirable, although the predominant consideration should generally be the

cost of service by the most efficient agency. In certain cases, especially before the Transportation Act of 1940 became effective, the Commission, in its zeal to conserve earnings, seems to have given too much weight to value factors and competitive relationships. Attempts to raise rates when lower charges can be shown to be reasonably compensatory, not only tend to encourage private transportation but also to prevent the movement of traffic according to the relative economy and fitness of the carriers. When business is inadequate in volume, it may be that all carriers cannot prosper unless minima are increased; yet it appears better to let the inefficient fall by the way than to shield high-cost transportation.

A general assumption that recent reductions in rates merely arise from competitive rate slashing and are not the result of more efficient methods of transportation seems unwarranted. It is, of course, true that rates have been cut below a reasonable cost level in many instances, and that the Commission must look to the long run as well as to the short run. However, a careful distinction should be made between the circumstances under which rates based upon fully-distributed costs are and are not defensible. Most dissenting opinions have arisen in connection with this point.[8] In some instances rates near out-of-pocket costs are desirable, provided these costs are properly defined. If out-of-pocket costs are identified with average variable costs, instead of marginal costs, when plant is approaching full utilization, rates based upon the former lead to an outright loss measured by the excess of marginal costs over variable. Possibly this is one reason for the Commission's reluctance to adopt the out-of-pocket cost standard.

The Commission's statement that minimum reasonable rates should be governed by the same standards applicable to maximum reasonable rates is misleading, except with reference to the general directives of Congress. It implies too much emphasis upon the value of the service as a basis for minima. Value has long been recognized as a leading test for maximum rates in industries like railroads, because it makes possible lower rates on low-grade traffic and thus a reduction in the total cost of service; but its use as a measure of minimum rates tends to raise the cost of service and to narrow the zone of discretion. If a distinction between maximum and minimum rates is to be recognized, it seems unreal-

8. Compare remarks by Commissioner Splawn in Petroleum Products between Kansas, Oklahoma, Arkansas, Missouri, and Colorado, op. cit., p. 638.

istic to base the latter upon the value of the service except to a limited extent on the part of a low-cost carrier characterized by constant costs. Today there is less room than formerly for the upward revision of rates according to what the traffic will bear, even in the railroad industry. The Commission itself has so stated.

Although some deviations have occurred, most minimum rate decisions of the Commission appear to be consistent with transportation by the low-cost agency. When rates proposed by that agency have been shown to be reasonably compensatory, i.e. cover the expenses which vary over a period of time with the traffic or use of the plant plus a proper return on investment, they have generally been approved. As in New Automobiles in Interstate Commerce, the Commission has often said that raising rates to protect uneconomic agencies is contrary to the national transportation policy and to the rules of rate making. Cases which have shielded high-cost carriers have been decided largely during a period of business depression and rapid technological change for the purpose of conserving carrier earnings, not out of favoritism, and are to be explained to some extent by the partial conflict in Congressional policies. Reconciliation of these policies under unfavorable conditions has called for a high degree of statesmanship on the part of the Commission.

<div align="right">TRUMAN C. BIGHAM.</div>

UNIVERSITY OF FLORIDA

THE MORTGAGE PORTFOLIO OF MUTUAL
SAVINGS BANKS

SUMMARY

I. Introduction, 232. — II. Supply and allocation of investment funds: deposits, 233; the three main portfolios, 237; the volume of mortgage lending, 245; factors influencing the allocation of funds, 248. — III. Mortgage interest rates, 256. — IV. Comparison with yields on other investments, 261. — V. Summary and conclusions, 264.

I. INTRODUCTION

The large volume of mortgage debt in the United States has created investment outlets for numerous financial institutions.[1] Among these the mutual savings banks are of prime importance, especially in the Northeast.[2] The 1940 Census revealed that, for the country as a whole, savings banks supplied 13.5 per cent of the total funds invested in outstanding mortgage loans on one- to four-family nonfarm houses; the same proportion obtained in 1942.[3] In some cities and regions the proportion is much higher. In Worcester, Massachusetts, the five mutual savings banks have traditionally enjoyed hegemony in the local mortgage market. The Financial Survey of Urban Housing revealed that in 1934 these banks held from two-thirds to three-fourths of all outstanding mortgages in Worcester, and the 1940 Census — based on an

1. In 1942 nearly $20 billion of mortgage loans was outstanding on one- to four-family nonfarm houses. FHLB Review, September 1943, p. 355. For various earlier estimates, see American Institute of Banking, Home Mortgage Lending (1938), pp. 7–9; J. H. Gray and G. W. Terborgh, First Mortgages in Urban Real Estate Finance (1929); C. E. Persons, "Credit Expansion, 1920 to 1929, and Its Lessons," this JOURNAL, November 1930, p. 104; Evans Clark and Associates, The Internal Debts of the United States (1933), p. 63; Charles Abrams, Revolution in Land (1939), pp. 98–99; D. C. Horton, Long Term Debts in the United States (1937); and J. W. Sternberg, "Indebtedness in the United States, 1929–39," Survey of Current Business, June 1940. See also American Housing: Problems and Prospects, Twentieth Century Fund (1944).

2. In 1929 the Brookings Institution estimated that mutual savings banks held $45 billion of mortgages, and in 1932 Evans Clark and Associates estimated that urban mortgages accounted for roughly 49 per cent of the assets of savings banks. Gray and Terborgh, op. cit.; Evans Clark and Associates, op. cit., pp. 63, 79.

3. U. S. Bureau of the Census, Census of Housing, 1940, Vol. IV, Part 1, p. 71; FHLB Review, September 1943, p. 355.

enumeration rather than a sample — corroborated this general figure.[4] The dominant rôle of the savings banks in the Worcester mortgage market is clear-cut, at least until the war. It therefore appears desirable to analyze in detail the mortgage portfolio of these banks, to shed light on the investment policies of an important group of mortgage lenders.

The nature of savings banks is too well known to require extended discussion here. A mutual savings bank has no capital stock; it is operated for the benefit of its depositors, who are the owners. It receives savings in the form of time deposits, which constitute roughly 90 per cent of total liabilities; the remaining liability items are the guaranty fund and profit and loss account, Christmas and other club deposits (segregated from the regular deposit accounts), interest and rents, less current expenses, and miscellaneous liabilities. It invests the aggregate funds chiefly in securities, real estate loans, and personal loans, and makes the profits available to the depositors in the form of "dividends" at regular intervals. Mutuals in Massachusetts and several other states may invest up to 70 per cent of total deposits in mortgage loans. Thus they not only supply credit for home financing, but in turn depend on the continuance of a sound mortgage market for investment yields. Since their deposits are not *legally* subject to withdrawal on demand, they can safely be invested in mortgages and other long-term loans.

II. Supply and Allocation of Investment Funds

The supply of mortgage credit available from savings banks depends on, but is not commensurate with, the volume of deposits. The proportion of deposits invested in mortgage loans is influenced by statutory and voluntary restrictions. The former sets the upper limit to the volume of mortgage lending. Within this statutory limit, the proportion of bank funds actually devoted to mortgages

4. On a sample of loans of all priorities in 1934, these banks held 70.8 per cent of mortgages on owner-occupied properties, and 71.3 per cent on rented properties. They held 76.8 per cent of all first mortgages. U. S. Bureau of Foreign and Domestic Commerce, Financial Survey of Urban Housing, 1934 (1937), p. 78. For the entire country, savings banks held only 12 per cent of total mortgages. Ibid., p. 76. In 1940 Worcester's savings banks held over 65 per cent of all first mortgages on one- to four-family owner-occupied houses. A breakdown reveals a ratio of almost 60 per cent on one-family houses, and over 73 per cent on two- to four-family houses. Census of Housing, 1940, Vol. IV, Part 2, p. 659. Percentages computed by the writer.

depends on investment policies and decisions which are based ultimately on the relative safety, liquidity, shiftability and yield of mortgage loans and alternative investments. We therefore proceed first to an analysis of deposits in the five Worcester savings banks from 1920 through 1940. The basic data were derived from the annual reports of the Commissioner of Banks.[5] The five mutual savings banks in Worcester were all well-established before 1920, the year in which our analysis begins.[6] Although they differ in size, their deposits and investments have followed a uniform pattern and have been subject to the same general influences. Their annual statements were therefore combined into aggregates for the five banks as a unit. In analyzing these figures, comparative data for all mutuals in the State will be used as a background.

Deposits. Table I summarizes the deposit experience of these banks. Aggregate deposits of the five banks stood close to $94 million in 1920, fell slightly during 1921, then rose rapidly and steadily to a peak of over $150 million in 1931. By 1933 they had fallen to $136.7 million, but recovered and, after a setback in 1938, reached over $146.5 million in 1940. The behavior of deposits is more clearly revealed by their annual increase or decrease. The cyclical pattern differs in timing from the general business cycle. The depression of 1921 and the recession of 1937–38 were both followed by a decrease in deposits, and business recovery brought an increase. But the prosperous year 1929 (October 1928 to October 1929)[7] witnessed a slackening in the rate of increase (only $3 million compared with over $10 million in 1928), whereas during the depressed years 1930 and 1931 the annual increase was greater than in 1929. This contradiction of expectations may be partly

5. Commonwealth of Massachusetts, Department of Banking and Insurance, Annual Report of the Commissioner of Banks. Public Document No. 8. Part I, Savings Banks and Institutions for Savings; Part II, Trust Companies, Foreign Banking Corporations and Others than Banks; Part III, Coöperative Banks and Savings and Loan Associations; Part IV, Credit Unions. Hereafter cited as Annual Report.

6. The banks, with their dates of incorporation and amount of resources in 1940, are as follows: Bay State Savings Bank, 1895, assets, $7,205,871; People's Savings Bank, 1864, assets, $34,918,813; Worcester County Institution for Savings, 1828, assets, $57,471,762; Worcester Five Cents Savings Bank, 1854, assets, $37,557,022; Worcester Mechanics Savings Bank, 1851, assets, $28,809,703. Annual Report, 1940, Part I.

7. Except where otherwise noted, the data in this study are not for the calendar year but for the fiscal year ending October 31. The discrepancy between the two accounting periods is slight.

TABLE I

DEPOSITS IN FIVE SAVINGS BANKS

Year	Total Deposits (000 omitted)	Annual Change (000 omitted)	No. of Accounts	Average Size
1920......	$93,857	....	191,664	$490
1921......	92,618	$1,241	188,229	492
1922......	95,521	2,905	190,581	501
1923......	103,489	7,968	199,717	518
1924......	109,626	6,137	205,025	535
1925......	116,105	6,479	209,969	553
1926......	121,345	5,240	212,731	570
1927......	129,297	7,952	215,096	601
1928......	139,570	10,273	219,014	637
1929......	142,647	3,077	218,584	653
1930......	146,566	3,919	216,592	677
1931......	150,335	3,769	213,710	703
1932......	140,945	−9,389	205,287	687
1933......	136,747	−4,198	202,108	677
1934......	138,641	1,894	203,130	683
1935......	140,448	1,807	203,568	690
1936......	143,139	2,691	206,392	694
1937......	145,355	2,216	209,465	694
1938......	142,834	−2,521	204,933	697
1939......	144,381	1,547	205,434	703
1940......	146,557	2,176	210,477	696

Source: Computed from 105 bank statements in Annual Report, Part I, 1920–40.

explained by the movements of the stock market. It is probable that large withdrawals for speculative purposes retarded the growth of deposits in 1929, and that after the stock market crash "scared money" and funds unable to find a more profitable outlet flowed into these banks.[8] Not until 1932 did the deposits of the Worcester mutuals reflect the widespread unemployment and industrial stagnation. By 1934 an increase in deposits was again evident.

The "average deposit" is a weighted arithmetic mean, computed by dividing aggregate deposits by the number of deposit

8. Cf. J. M. Keynes: "Now when bullish sentiment is on the increase there will be a tendency for the savings deposits to fall." Treatise on Money, Vol. II, pp. 250–251. On the other hand, savings deposits may be withdrawn on a declining stock market in order to meet margin requirements. For discussion of the cyclical activity of savings accounts in New York State, see W. Welfling, "Some Characteristics of Savings Deposits," American Economic Review, XXX, December 1940, pp. 748–758.

accounts in the five banks. The average deposit increased steadily
to 1931, fell only slightly in 1932 and 1933, and then rose. Per
capita deposits were derived by dividing the aggregate deposits of
the five banks by the population of Worcester. Per capita deposits
increased to 1930, fell slightly during the depression, and then
increased. The number of deposit accounts grew steadily after 1922
and reached an all-time peak of 219,014 in 1928, then fell to a low
point in 1933. By 1940 there were 210,477 accounts. In each year
of our study the number of accounts exceeded the population of
Worcester. In 1920 there were 107 accounts for every 100 persons;
in 1940 there were 109. These figures indicate the popularity and
stability of the mutuals in Worcester.[9] But the high ratio of
deposit accounts to population should be interpreted conserva-
tively. All accounts are not owned by residents of Worcester;
many individuals have accounts in several banks; and many
families open an account for each child. Some accounts, moreover,
represent deposits of churches and fraternal groups rather than of
individual savers, and many represent small school savings, which
were popularized and encouraged during the 'twenties.[1] Neverthe-
less, the figures do indicate the widespread character of savings in
Worcester, and the large number of people who are the ultimate
sources of mortgage credit.[2]

9. In Baltimore, where the mutuals have also had a long history, the
number of deposit accounts in 1934 equalled only 40 per cent of the population,
or one account for every two and one-half persons. R. W. Thon, Jr., Mutual
Savings Banks in Baltimore (1935). The comparison should not be stressed,
however, since Baltimore is a larger city than Worcester, and savings banking
is not so well entrenched in Maryland as it is in Massachusetts.

1. For example, during the school year 1926–27, in Worcester and the
adjacent town of Auburn, 14,307 pupils (55 per cent of all pupils enrolled in
60 schools) deposited $36,339, or $2.54 per depositor. An average of 4,532
pupils made deposits each week. American Bankers Association, Savings
Bank Division, School Savings Banking During School Year 1926–27 (1927),
p. 14.

2. For all savings banks in the State, the behavior of deposits was similar
to that in Worcester, with the following variations: there was no decline in
total deposits in 1921; the number of accounts reached a peak in 1929 rather
than in 1928, and a trough in 1934 rather than in 1933, and fell slightly in 1936,
1938 and 1940. The average deposit account in Worcester is slightly smaller
than in the State as a whole. Annual Report, 1940, Part I, p. xxiv However,
per capita deposits in the Worcester mutuals have been higher than per capita
savings of all types in the State. The latter increased from $384 in 1921 to
$670 in 1931, declined during 1932 and 1933, then recovered and reached
$538 in 1936. During the entire period since 1920, Massachusetts continued
to lead the country in per capita savings. American Bankers Association

The five banks have enjoyed considerable secular growth in savings, the 1940 figure being greater than that for 1920.[3] Their deposits increased rapidly during the 1920's, and again since 1934, in the face of generally declining dividend rates. Since they may legally invest up to 70 per cent of deposits in mortgage loans, a considerable amount of mortgage credit is potentially available from these banks. Their deposits are derived from a large number of small depositors and their long-run stability seems assured. This would indicate a growing supply of potential mortgage money at cheaper cost to the savings banks, and hence the possibility of lower mortgage rates to borrowers. However, deposits have been affected by the business cycle, and in some years a sharp contraction in savings has reduced the amount of potential mortgage funds.

The Three Main Portfolios. Table II presents aggregate assets of the five banks and the totals of the three main portfolios: mortgage loans, securities, and personal loans. These three portfolios are also calculated as percentages of total assets in each year, and the resulting ratios are plotted in Chart I. In this section the chief emphasis is on securities, which are the main alternative to mortgage loans. The latter are briefly compared with the other two portfolios, but a more detailed analysis of the volume of mortgage lending is deferred until later. The purpose of the statistical comparison of investment portfolios is to show the actual relationship between mortgage loans and other investments during the period studied; their relative yields and the factors influencing the allocation of funds will be examined below.

Total assets exhibit the same pattern as total deposits, although the aggregate amounts are of course larger. Mortgage loans grew to a peak of nearly $90.5 million in 1929, declined slightly through 1931, and fell every year thereafter. The securities portfolio grew to over $64.7 million in 1931, declined during the next two years, then increased to over $83.8 million in 1940. Personal loans fluc-

Savings Division, Savings Deposits and Depositors in Banks and Trust Companies in the United States (1921, 1926, 1929, 1930, 1931, 1934, 1936 and 1937).

3. The growth of savings deposits during the 1920's resulted partly from a general shifting of deposits from the demand to the time category. Depositors desired the higher interest rates received on time deposits, while commercial banks wished to take advantage of the three per cent reserve permitted by the Federal Reserve System against time deposits. W. H. Steiner, "Savings Banks," Encyclopedia of Social Sciences, Vol. XIII, pp. 556.

Year	Total Assets	Mortgages	Securities	Personal Loans
1920.........	$102,291	$51,748	$35,187	$12,751
1921.........	101,655	54,147	35,086	9,992
1922.........	105,431	57,175	38,486	7,755
1923.........	114,562	63,086	40,946	8,396
1924.........	121,807	71,775	40,851	6,551
1925.........	129,392	79,578	41,263	6,202
1926.........	135,641	84,340	42,070	7,010
1927.........	144,509	86,955	47,592	6,844
1928.........	155,921	89,370	53,751	9,806
1929.........	160,619	90,472	54,992	12,809
1930.........	165,144	90,447	60,145	10,182
1931.........	168,830	90,269	64,745	8,121
1932.........	158,462	87,934	55,758	3,675
1933.........	153,397	85,401	54,491	2,558
1934.........	154,956	81,620	58,918	1,877
1935.........	156,797	78,978	62,988	1,581
1936.........	160,026	77,591	69,127	1,246
1937.........	163,314	76,799	76,040 ·	767
1938.........	161,244	75,700	75,643	670
1939.........	163,510	74,362	78,764	669
1940.........	165,963	73,482	83,804	572

Source: Computed from 105 bank statements in Annual Report, Part I, 1920–40.

tuated during the 'twenties, reached a peak in 1929, and fell sharply during the 'thirties.

The relationship among alternative investments is revealed by computing each as a percentage of total assets (Chart I). The ratio of mortgage loans to assets increased from 50.6 per cent in 1920 to a peak of 62.2 per cent in 1926, declined to 53.5 per cent in 1931, increased to 55.7 per cent in 1933, and fell to 44.3 per cent in 1940. In 1920 the five banks held 12.5 per cent of total assets in personal loans. Since then, except for an increase in 1928 and 1929, personal loans have shrunk in absolute and in relative terms. From 8.0 per cent in 1929 they declined to 0.34 per cent of assets in 1940.

As Chart I shows, the securities-assets ratio exhibits a pattern almost directly opposite that of mortgages to assets. From 1923
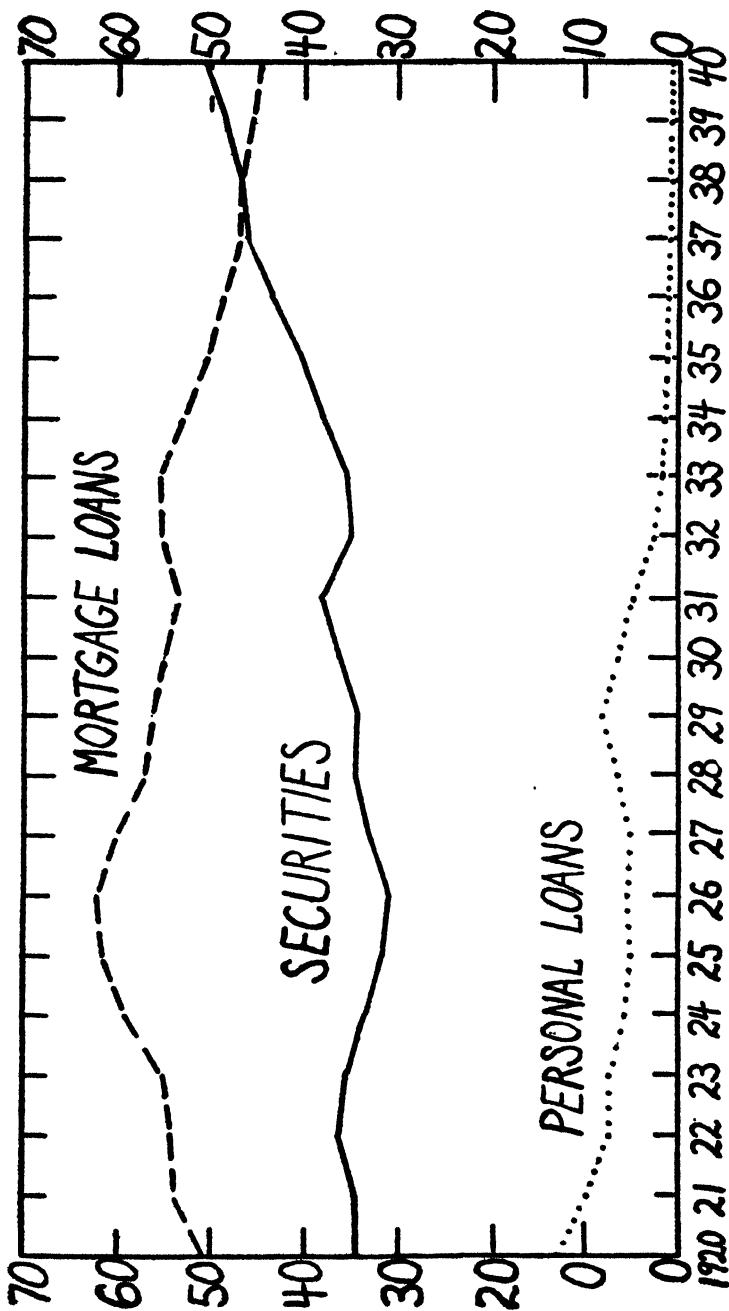
CHART I
INVESTMENTS AS PER CENT OF ASSETS

through 1926, when the mortgage-asset ratio was *rising*, the securities-asset ratio *fell*. This suggests that as the real estate market created a greater demand for mortgage loans, funds were withdrawn (relatively, not absolutely) from the securities portfolio and were placed in mortgage loans. From 1927 through 1931, when the mortgage-asset ratio was falling, the securities-asset ratio increased, with only a slight drop in 1929, although even in that year the absolute increase in security holdings continued. During 1932, when the mortgage ratio increased, the securities ratio fell. In absolute amounts, assets, mortgage loans and securities all declined during 1932, but the fall in mortgages was smaller than the decline in assets, resulting in a higher ratio. Whereas mortgages declined by only about $2.33 million in 1932, securities fell by nine million dollars. The probable explanation is that mortgage investments are relatively inflexible in time of depression, and a bank must therefore sell some of its highly marketable securities in order to provide cash to meet withdrawals. During 1933, as assets declined faster than either mortgages or securities, both ratios rose. After 1933 the movements of mortgages and of securities have been diametrically opposed, both in absolute and in relative terms. Mortgages fell from $85.4 million, or 55.7 per cent of assets in 1933, to less than $73.5 million, or 44.3 per cent of assets in 1940. Securities, on the other hand, grew from less than $54.5 million, or 35.5 per cent of assets in 1933, to over $83.8 million, or 50.5 per cent of assets in 1940. At the end of the fiscal year 1938, securities and mortgages each accounted for 46.9 per cent of assets, and by 1939 securities had outstripped mortgages.

Since the funds of the five mutuals have been increasingly invested in securities, as mortgage loans have declined, we next examine the composition of the securities portfolio. Table III and Chart II present a breakdown of this portfolio into classes of security, by amount held in selected years and as a percentage of each of total assets.

"Public funds, bonds and notes" were unclassified until 1934, when United States Government securities and fully guaranteed obligations were distinguished from "other public funds, bonds and notes," the latter category including securities of states, cities and towns, and special districts. Of the total assets of the five Worcester mutuals, public securities constituted 15.4 per cent in 1923, only 10.8 per cent in 1927, and 11.2 per cent in 1931. Since

TABLE III

SECURITIES PORTFOLIO, FIVE SAVINGS BANKS, SELECTED YEARS
(In thousands of dollars)

Securities	1923	1927	1931	1936	1940
Total..........	40,946	47,592	64,745	69,127	83,804
U.S. Government ⎫	17,677	15,615	18,903	⎰ 36,636	54,341
Other Public.... ⎭				⎱ 6,081	7,894
Railroad.........	17,416	19,014	24,827	14,451	8,918
Street Railway....	2,518	2,182	1,550	792	255
Telephone........	1,861	2,526	3,319	1,148	2,107
Utilities..........	604	6,648	11,727	7,014	7,379
Bank Stock.......	720	1,289	3,835	2,666	2,769
Other* ...¦.......	132	319	584	339	142

* Boston Terminal Co. bonds, Securities Acquired for Debts, and Federal Land Bank bonds

Source: Computed from 25 bank statements in Annual Report, Part I, 1923, 1927, 1931, 1936 and 1940.

then, these investments have grown rapidly. By 1940 public securities of all kinds had increased to 37.5 per cent of assets (32.7 per cent federal, and 4.8 per cent other public securities). Railroad bonds and notes are the second largest category. From 15.2 per cent of assets in 1923 they declined to 13.2 per cent in 1927, but increased to 14.7 per cent in 1931. During the 'thirties they declined both absolutely and relatively, falling to 5.4 per cent of assets in 1940. Street railway bonds declined both absolutely and relatively during both decades. In 1923 they constituted 2.2 per cent of assets, but by 1940 they were only 0.15 per cent. Boston Terminal Company bonds were always a very small item.

Telephone company and public utility bonds increased rapidly during the latter 'twenties, utilities experiencing the more rapid growth. During the 'thirties both suffered an absolute as well as a relative decrease. Telephone company bonds increased from 1.6 per cent of assets in 1923 to 2.0 per cent in 1931. By 1936 the ratio had fallen to 0.7 per cent, and it remained at that level through 1939. In 1940 these securities doubled in absolute amount and the ratio increased to 1.26 per cent. Gas, electric and water company bonds increased from 0.5 per cent of assets in 1923 to 6.9 per cent in 1931. By 1936 this ratio had fallen to 4.4 per cent, and by 1939 it was only 2.9 per cent. A substantial absolute increase raised this ratio to 4.4 per cent in 1940. Investment of the Worcester
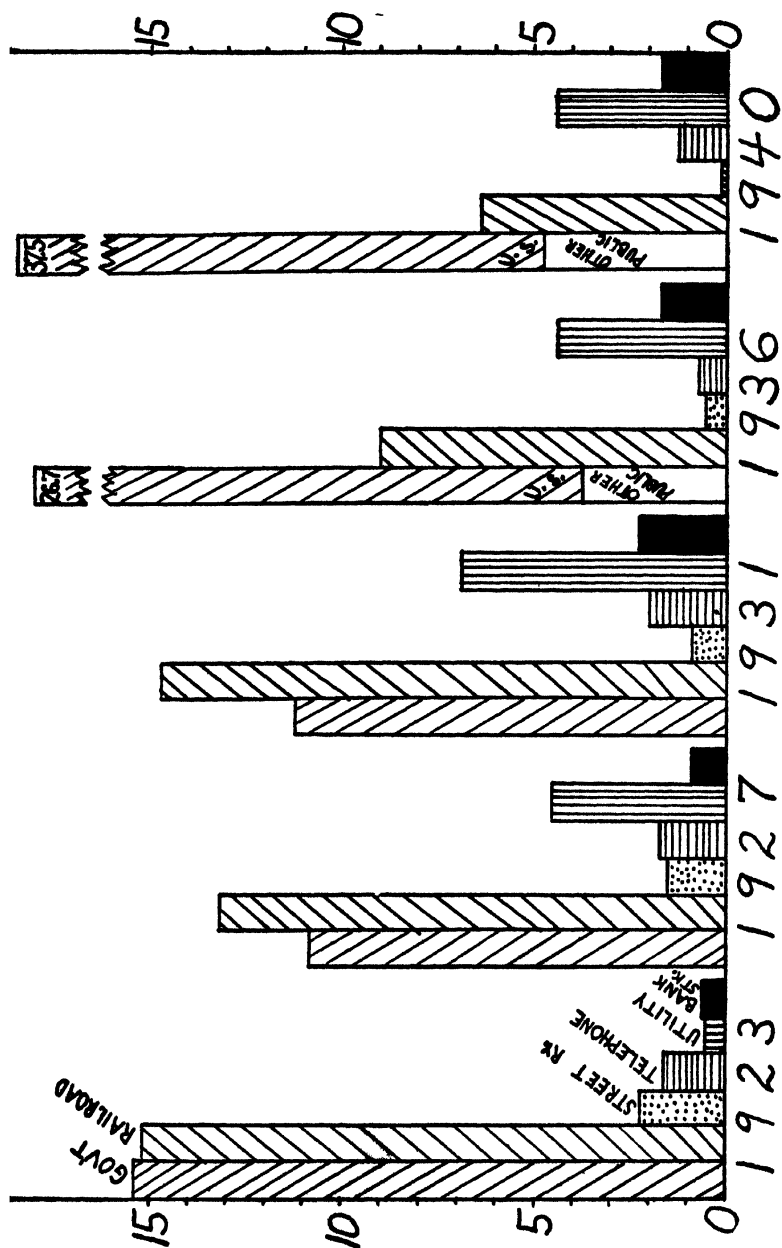
CHART II

VARIOUS SECURITIES AS PER CENT OF ASSETS

mutuals in bank and trust company stocks increased both relatively and absolutely during the 'twenties, but decreased during the 'thirties. In 1923 they constituted 0.6 per cent of assets; in 1931, 2.3 per cent. By 1936 the ratio had fallen to 1.7 per cent, and it remained at this level in 1940. In the Worcester banks (and in all Massachusetts mutuals), miscellaneous items such as bankers' acceptances, federal land bank stocks, and securities acquired for indebtedness comprised a negligible portion of total security holdings and of total assets.[4]

Assets other than the three investment portfolios discussed above have usually been small. Massachusetts mutuals have always got along with a relatively small amount of available cash, some of it held in the Mutual Savings Central Fund and earning interest, but most of it deposited in other banks and not on interest. From 1920 through 1931 cash and cash items averaged around 1.7 per cent of total assets. In 1932 and 1933 the necessity for keeping in a liquid position raised this ratio to 2.51 per cent and 2.38 per cent, respectively. In 1938 the state-wide ratio of cash

4. In all mutuals in the State, the experience described in this section was roughly similar. The mortgage-asset ratio rose from 43.9 per cent in 1920 to a peak of 54.3 per cent in 1926, remained fairly stable at slightly over 53 per cent through 1933, and thereafter declined to 40 per cent in 1940. (It will be noted that mortgage loans in the five Worcester mutuals have constituted a larger proportion of assets than in all mutuals in Massachusetts.) From almost 10 per cent in 1923, the personal loan-asset ratio declined during the decade, while the mortgage portfolio was expanding. In 1929 the ratio increased, temporarily, to 9.7 per cent. By 1940 it had declined to 0.89 per cent. (The Worcester banks have usually held a smaller proportion of assets in personal loans than have all the mutuals in the State.) In 1924 the state-wide securities-asset ratio was 36.3 per cent; in 1929 it was 34.3 per cent. It increased to 38.7 per cent in 1934, 47.5 per cent in 1939, and 48.4 per cent in 1940.

The decline in government security holdings during the 'twenties and the great increase during the 'thirties took place throughout the State. In 1923, 20 per cent of assets were held in these securities, but by 1930 the ratio had fallen to 11.1 per cent. The ratio had increased to 17 per cent by 1934 and to 33.1 per cent by 1940 (29.1 per cent federal and 4.0 per cent other public securities). Railroad and street railway bonds declined during the two decades. In 1925 railroad bonds were 12.9 per cent and street railway bonds were 2.1 per cent of assets. By 1940 these investments had fallen to 7.8 per cent and 0.3 per cent of assets, respectively. Boston Terminal Company bonds were insignificant, as was the case in the Worcester mutuals. Throughout the State, telephone and public utility company bonds both enjoyed an increase to about 1933, the latter growing more rapidly than the former. After 1933 these investments declined relative to total assets. Finally, in all Massachusetts mutuals, holdings of bank stocks fluctuated during the early years of the 1920's, but in general they increased from 1.07 per cent of assets in 1920 to 2.14 per cent in 1940. Annual Report(s), Part I, 1920–40.

to assets was 2.75 per cent; by 1940 it had increased to 4.07 per cent. Real estate for banking purposes, furniture and fixtures, and sundry assets together constituted only 1.07 per cent of total assets of Massachusetts mutuals in 1940. The remaining asset item is real estate held by foreclosure, which was a fraction of one per cent of assets during the 'twenties, increased to some 8 per cent during the depression years, and was reduced to about 5 per cent by 1940. These ratios were roughly the same for the five Worcester banks.[5]

In summary, from 1923 through 1926 the banks directed a growing proportion of their funds into mortgages. Part of these funds came from a proportionate reduction of investment in public securities, railroad bonds and notes, and street railway bonds. As the demand for mortgage money began to decline, the banks invested a larger portion of their funds in securities, the legal list of securities having been broadened to create new investment outlets (see below, p. 254). During the latter 1920's this took the form of increased holdings of public utility and telephone company bonds, which together absorbed almost 10 per cent of total assets by 1931. During the depression these securities suffered in yield, and every type of security except governments declined in absolute amount and as a percentage of assets from 1931 to 1939. The small increases in public utility bonds, telephone company bonds and bank stocks during 1940 were insufficient by themselves to absorb the growing deposits and the funds released from the partial liquidation of other securities. Thus, during the 1930's, the growing funds which could not find an outlet in the mortgage market were invested in the greatly expanded federal debt.

Mortgage loans have a higher yield than almost any other bank investment (see Section IV, below). It is therefore reasonable to assume that funds will be placed in mortgages, if possible. This was the case to a large extent. The growing mortgage-asset ratio up to 1926 and the falling ratio after 1926 coincide with the rise and fall of the real estate cycle, and hence the allocation of funds between mortgages and other investments seems to depend in great measure on the *demand* for mortgage credit. Whenever the demand for mortgage loans has declined, the securities portfolio has expanded. A certain portion of bank funds *should* be invested in securities, in order to maintain shiftability; but certainly the

5. Annual Report(s), Part I, 1920–40; and Annual Report, 1940, p. xi.

need for maintaining shiftability does not require the investment of over half the total assets in securities, especially low-yielding government securities, as these banks have done. However, funds so invested do earn something, and the heavy portfolio of government securities can easily be liquidated, if and when the demand for mortgage loans should increase. The inverse relationship of the mortgage and the securities portfolios, and the close relationship of the mortgage portfolio to the real estate cycle, suggest that the expansion and contraction of the securities portfolio is not a cause but a consequence of the changes in mortgage-lending volume. The allocation of investment funds is not due solely to changes in demand for mortgages, however. As pointed out below, various legal and economic factors condition mortgage loans and alternative investments.

The Volume of Mortgage Lending. The number and value of (new and refinanced) real estate loans made annually, and the total loans held by each bank were first derived from the annual reports published in the bank commissioner's reports, and were
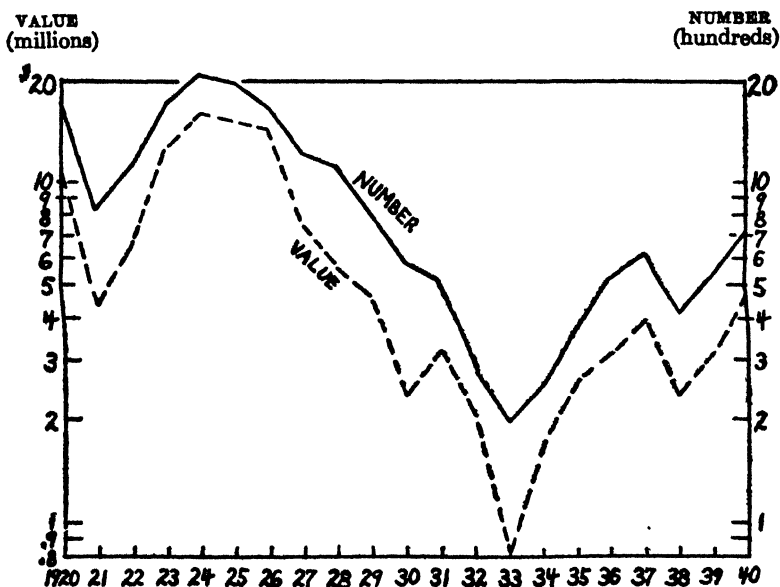
CHART III
New Loans Annually
(Includes Refinancing)
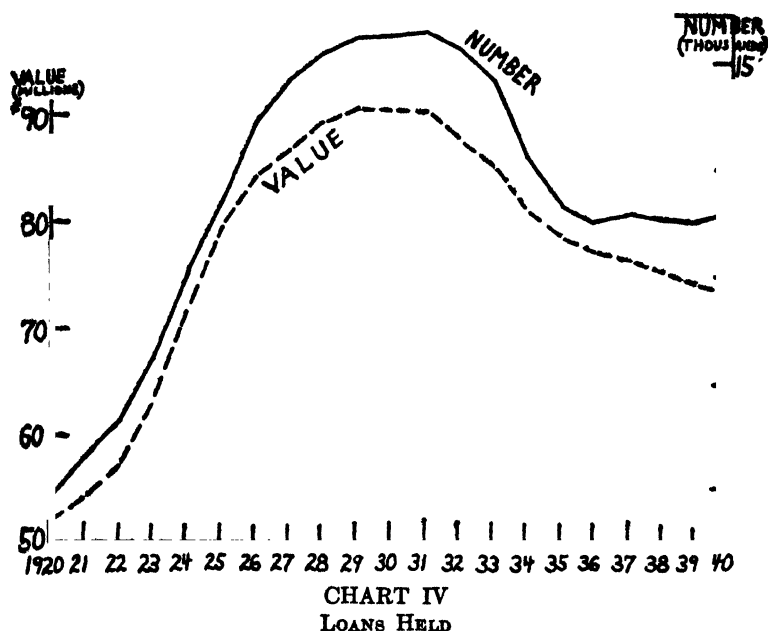
CHART IV
LOANS HELD

then combined by the writer into aggregates for the five banks.
Table IV and Charts III and IV summarize the results.

The *number* of mortgage loans held in the combined portfolio
of the five banks (Chart IV) grew steadily to a peak in 1931, then
fell and reached a trough in 1936, and finally increased slightly
by 1940. These figures are for the number of loans held at the end
of October of each year. A more significant series is the number of
new or refinanced loans made during each fiscal year. This shows
fluctuations in mortgage lending more clearly, because even a large
increase in new loans may represent but a small increase in the
total number of loans held in the portfolio. Except for a sharp
drop in 1921 and 1938, Chart III shows a rise in the number of
loans made from 1920 to a peak in 1924, a decline from 1925 to a
trough in 1933, and a rise thereafter. Although the *rate* of new
lending fell off after 1925, the net addition to the mortgage port-
folio continued through 1931; and whereas the rate of new lending
increased from 1934 through 1937, the decrease in total number
of loans held (which began in 1932) continued through 1936. The
obvious explanation of this phenomenon is that from 1925 through
1931 the number of new loans made annually exceeded the number

of discharges, while from 1932 through 1936 the number of discharges was greater than the volume of new loans made annually. The number of new loans made conforms to the real estate and building cycle, while the number of mortgages held bears a closer resemblance to the general business cycle.

With one exception, the *value* of loans made annually moved in the same direction as the *number* of loans made (see Chart III). The exception was in 1931, when the value of loans made was $852,700 *greater* than in the previous year, while the number of loans made was 70 *less* than in the previous year. A greater average size of new loans, or the voluntary or involuntary extinction of

TABLE IV

MORTGAGE-LENDING DATA, FIVE SAVINGS BANKS
(Value in thousands of dollars)

Year	New Loans Annually*		Loans Held		
	Number	Value	Number	Value†	Annual Change
1920 ...	1,665	$10,769	10,921	$51,748	.....
1921 ...	818	4,261	11,300	54,147	$2,399
1922 ...	1,089	6,406	11,674	57,175	3,027
1923 ...	1,660	12,082	12,210	63,086	5,911
1924 ...	2,028	15,882	13,027	71,775	8,689
1925 ...	1,959	14,638	13,715	79,578	7,803
1926 ...	1,618	12,060	14,449	84,340	4,762
1927 ...	1,215	7,769	14,862	86,955	2,614
1928 ...	1,121	5,727	15,102	89,370	2,415
1929 ...	805	4,617	15,239	90,472	1,102
1930 ...	587	2,319	15,282	90,447	− 24
1931 ...	517	3,171	15,325	90,269	− 178
1932 ...	274	1,076	15,158	87,934	−2,336
1933 ...	197	819	14,831	85,401	−2,532
1934 ...	254	1,680	14,129	81,620	−3,782
1935 ...	374	2,611	13,678	78,978	−2,642
1936 ...	521	3,169	13,509	77,591	−1,387
1937 ...	618	3,944	13,598	76,799	− 793
1938 ...	418	2,361	13,531	75,700	−1,099
1939 ...	527	3,193	13,504	74,362	−1,338
1940 ...	719	4,657	13,612	73,482	− 880

* Including refinancing.
† Excluding amounts due to mortgagors.
Source: Computed from 105 bank statements in Annual Report, Part I, 1920–40.

many small loans, may account for the above situation. In the case of loans held (Chart IV), there are a few slight discrepancies in direction of movement between the number and the value data. Although the *number* of loans held increased somewhat in 1930, 1931, 1937 and 1940, the *value* of loans held declined slightly in each of those years. In all other years the number and value data exhibit the same pattern.[6]

In summary, the mortgage portfolio of the five banks was considerably expanded during the real estate boom of the 'twenties, reached a peak in 1929, and was maintained at a fairly stable level through 1931. Since then it has shrunk, measured by the amount of mortgages held each year. The number and value of (new and refinanced) loans written each year has been increasing since 1934, however, and it is possible that at the start of the war the mutuals were at the turning point and were beginning another expansion of their mortgage portfolio. Measured by volume of new loans made, the lending pattern coincides with the real estate cycle rather than the general business cycle. There is also evidence of a sharp secular decline, both in number and value, since 1924. The number of loans made (or refinanced) during 1939 was only 26.0 per cent of the number made in the peak year 1924. This ratio increased to 35.5 per cent in 1940. The dollar value of loans made during 1939 was only 20.1 per cent of the amount made during 1924, but this ratio increased to 29.3 per cent in 1940. To a great extent this decline follows the real estate cycle, but this is not a complete explanation, for there was a rise in real estate activity during the latter 1930's.[7]

Factors Influencing the Allocation of Funds. The allocation of investment funds and the volume of mortgage lending, analyzed

6. The mortgage-lending experience of all mutuals in the State was roughly similar to that of Worcester's banks, but with the following variations: in number of loans held, a low point was reached in 1937 rather than in 1936; the annual number of loans made declined only slightly in 1921, reached a trough in 1934 rather than in 1933, and showed no decline in 1938. Annual Report(s), Part I, 1920–40. In dollar amount, the annual increase or decrease in the state-wide mortgage portfolio also differed somewhat from Worcester's banks. Annual Report, Part I, 1940, p. xxiv.

7. For the nation as a whole, the only significant shift in the proportionate amount of mortgage business (loans made annually on one- to four-family houses) between the late 'twenties and the late 'thirties was a decline by the mutual savings banks from 14 per cent in 1926–29 to 4 per cent of the total in 1937–40. Miles L. Colean, American Housing, op. cit., pp. 252–253.

above, tell little concerning the forces which affect the supply of available mortgage money. These conditioning factors acting upon the primary supply factor (the supply of savings) may be classified as legal restrictions and economic factors affecting investment policy. We shall discuss them in this order.

Legal restrictions on mortgage investment. The chief legal restrictions on mortgage investment concern the ratio of the total mortgage portfolio to the total deposits of the lending institution, and the ratio of each loan to the appraised value of the property. Since 1876 the mutuals of Massachusetts have been permitted to invest not more than 70 per cent of total deposits in first mortgage loans.[8] After 1908 the same restriction applied to trust company savings departments.[9] Therefore, in the period 1920–40, the statutory limit to the mortgage funds available from mutuals and trust company savings departments (exclusive of selling or borrowing upon mortgage paper) was 70 per cent of total deposits. Throughout most of this period, however, the actual ratios were below the legal maximum, as Table V and Chart V show.

The mortgage-deposit ratio in the five mutuals rose to 69.5 per cent in 1926, fell to 60 per cent in 1931, rose slightly in 1932 and 1933, and then fell to 50.1 per cent in 1940. From about 1924 to 1929 there was evidently a strong pressure of demand upon the legally available supply of mortgage funds, the ratio being over 63 per cent in each of those years. As the statutory limit was approached in 1926, the banks probably began to discourage some borrowers and to ration their funds among the better credit risks. At the same time a greater proportion of funds was directed into securities. In most years since 1920 there were only one or two trust company savings departments in Worcester. In view of their high mortality, too much weight should not be placed upon their ratios. Certain significant facts are nevertheless apparent. From 1923 through 1932 the ratio remained over 54 per cent (with a peak of 67 per cent in 1926), indicating the pressure of demand for real estate loans. The sharp drop in 1934 was caused by the

8. This legal ratio had previously been 75 per cent. In Rhode Island the legal ratio is also 70 per cent of deposits; in Maine, 60 per cent; New Hampshire, 75 per cent; and New Jersey, 80 per cent. Connecticut's mutuals may lend up to 70 per cent of deposits and surplus; Minnesota, 70 per cent of total assets; and Vermont, 80 per cent of total assets. American Institute of Banking, *Home Mortgage Lending*, p. 96.

9. Commonwealth of Massachusetts, Acts of 1908, Chap. 520.

TABLE V

MORTGAGE-DEPOSIT OR MORTGAGE-ASSET RATIO, SELECTED INSTITUTIONS

Year	Mutuals*	Trust Companies		Coöps. and F. S. and L. A.†	Credit Unions†
		Savings Dept.*	Other Dept.†		
1920 ...	55.1%	27.9%	7.5%	91.0%	....
1921 ...	58.5	32.0	8.9	92.6	....
1922 ...	59.9	34.4 .	8.1	89.8	....
1923 ...	61.0	59.9	10.3	90.6	....
1924 ...	65.5	63.2	11.0	87.0	0.2%
1925 ...	68.5	61.4	11.8	92.9	6.4
1926 ...	69.5	67.0	12.2	94.7	52.0
1927 ...	67.3	66.4	12.1	91.4	57.0
1928 ...	64.0	61.8	13.0	92.4	48 0
1929 ...	63.4	56.5	12.9	93.3	41.2
1930 ...	61.7	58.3	11.2	90.6	46.9
1931 ...	60.0	54.3	10.9	89.1	47.1
1932 ...	62.4	56.7	11.4	86.2	47.1
1933 ...	62.5	48.7	14.9	81.4	46.4
1934 ...	58.9	6.7	8.7	75.5	35.0
1935 ...	56.2	10.1	6.6	69.6	24.4
1936 ...	54.2	10.8	5.4	72.3	22.8
1937 ...	52.8	12.4	4.6	‡	19.4
1938 ...	53.0	15.1	4.0	‡	19.9
1939 ...	51.5	17.7	3.3	90.3	28.6
1940 ...	50.1	22.1	2.7	91.3	29.1

* Mortgage-deposit ratio.
† Mortgage-asset ratio.
‡ No data.
Source: Computed from 334 bank statements in Annual Report, Parts I and II, 1920–40; Part III, 1920–36; Part IV, 1924–40; and Annual Statement of the Worcester Cooperative Federal Savings and Loan Association, December 31, 1940.

establishment of a large savings department in a newly reorganized trust company which, however, had only a very small mortgage portfolio. Since 1934 the low (though rising) ratio in the trust company savings departments, and the falling ratio in the mutuals, seem to indicate that other factors than legal restrictions (lack of demand, competition of other lending agencies, etc.) have reduced the proportion of deposits invested in mortgages by these banks. Thus we consider briefly those institutions whose mortgage portfolios were *not* restricted by law.

Coöperative banks in Massachusetts were designed to encourage home ownership as well as thrift; their funds have traditionally
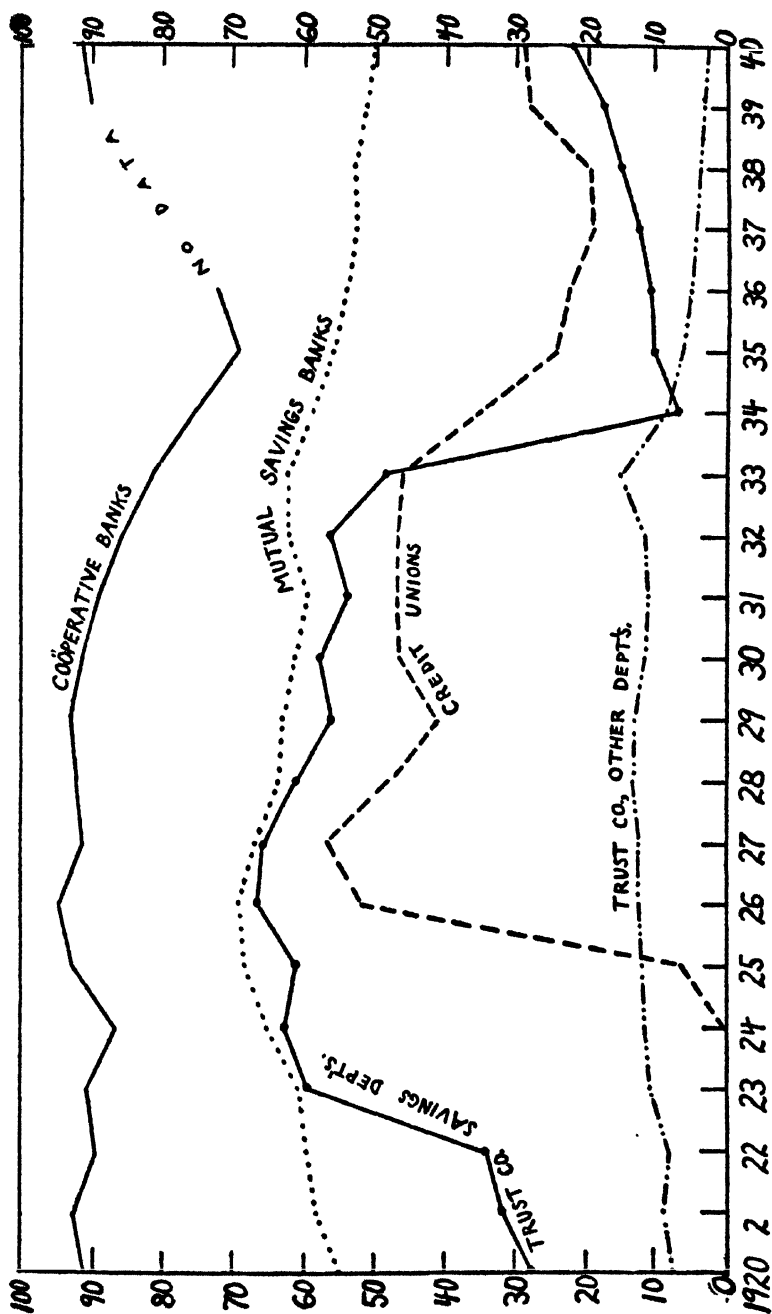
CHART V

MORTGAGE LOANS AS PER CENT OF DEPOSITS OR ASSETS

been invested almost entirely in mortgages. Nevertheless, the mortgage-asset ratio has fluctuated. It was over 90 per cent during most of the 'twenties (94.7 per cent in 1926), then fell to 69.6 per cent in 1935. After federalization and consolidation of the four coöperative banks into one savings and loan association in 1937, the ratio rose rapidly, reaching 91.3 per cent in 1940. A roughly similar pattern is found in trust companies and credit unions. Although trust officers enjoy wide discretion in investing funds of the commercial and trust departments, such funds are typically not invested in real estate loans. The mortgage-asset ratio rose slowly to 1928, then fell (except in 1932 and 1933) to 2.66 per cent in 1940. Finally, in those credit unions which held mortgages, the ratio of the latter to their assets reached 57 per cent in 1927, fell to 1929, remained over 46 per cent from 1930 through 1933, declined to 19.4 per cent in 1937 and rose to 29.1 per cent in 1940. In brief, all lending institutions examined exhibit a roughly similar mortgage investment pattern, *whether limited by law or not.*

It therefore appears that the legal restriction was not the sole — or even the main — factor which prevented the mortgage-deposit ratio in mutuals and in trust company savings departments from exceeding 70 per cent in the early 'twenties and brought about a reduction in that ratio after 1926.[1] Probably the chief factor in the allocation of savings bank funds into mortgages and other channels is the demand for mortgage loans, which, as indicated by the real estate cycle, has corresponded fairly closely to the volume of mortgage lending. When the demand is so great as to raise the mortgage-deposit ratio close to the legal maximum, the statutory

1. Although the statutory limit on mortgage investment was not reached in most years, it nevertheless exerts some influence on investment policy *before* that point. During a depression, deposits usually shrink faster than the mortgage portfolio can be reduced, since mortgage loans are long-term contracts and considerations of "goodwill" and the need of some borrowers to renew or extend matured loans makes difficult a sudden shrinkage in such loans. Under such conditions the mortgage-deposit ratio rises, as it did in the mutuals in 1932, 1933 and 1938, and in trust company savings departments in 1930 and 1932. Should a bank be loaned up to the statutory limit of 70 per cent, a contraction of deposits might cause a higher, and illegal, ratio. Hence the necessity for a "cushion" of about 10 or 15 per cent makes the "normal" ratio of mortgages to deposits about 55 or 60 per cent. The legal restriction on mortgage investment therefore affects the mortgage portfolio directly through the 70 per cent maximum ratio, and also indirectly by influencing sound investment policy to take into consideration cyclical changes in the volume of deposits.

restriction begins to operate; in times of slack demand the legal restriction has little or no effect.[2]

Another legal restriction is the maximum debt-value ratio on individual loans. Until 1935 mutual savings banks could lend only 60 per cent of the appraised value of the property on a first mortgage. This ratio has since been raised to 80 per cent, and on properties valued under $6,000 it may be 90 per cent on FHA insured loans (100 per cent on some G. I. loans). In other types of lending institutions there has been a similar trend toward higher debt-value ratios. With a given demand for housing, and a given number of mortgage loans to finance home ownership and residential construction, a bank will be in a position to invest more funds in mortgages, the larger the amount it may lend on each loan. The 60 per cent limitation operated as a legal restriction on the mortgage portfolio during the real estate boom of the 'twenties, when borrowers desired larger loans and wished to invest a smaller proportion of their own cash. The higher ratios now permissible increase the supply of mortgage funds legally available. More important, it increases the *demand* for first mortgage loans, but decreases the demand for junior financing. Finally, the higher debt-value ratios prevailing in the past ten years have influenced bankers to scrutinir ᴜhe credit standing of borrowers more carefully, and to place less reliance on a conservative debt-value ratio as a protection to the lending institution.

Factors influencing alternative investments. Other things equal, investment in mortgages will be restricted when attractive alternative investment channels exist, and will be expanded when legal or economic factors narrow the field of investment in other than mortgage loans. For mutual savings banks in Massachusetts, the element of safety is achieved both by legal restrictions and by investment traditions. Investment is limited to specified types of securities, and a limit is placed on the proportion of funds which may be invested in any single class of securities. No single loan or investment may exceed five per cent of the bank's deposits and income.

In 1908 investments in telephone company bonds were legalized for mutuals, but for many years the banks could invest only in the bonds of a Massachusetts corporation (the American Tele-

2. The demand for mortgage loans is too complex to be treated summarily here. It is the subject of a forthcoming article by the writer.

phone and Telegraph Company). The provisions of this law were broadened in 1925 and 1928. When public utility bonds were first legalized in 1919, the mutuals were restricted to Massachusetts companies. In 1926 this provision was broadened to include companies outside the State. This broadening of the field of investment gave new outlets to many savings banks, then approaching the 70 per cent limit on mortgage loans. As a result, public utility bondholdings in savings banks increased from 1926 to 1933, both in the State and in Worcester (see Table III, above). The statutory restriction of public utility investments to 15 per cent and of telephone company bonds to five per cent of deposits has probably been an element of safety, and has prevented an undue limitation upon the mortgage portfolio in those years when the high yield of such bonds made them appear relatively attractive. Their declining yield during the depression reduced their ability to compete with mortgages in the allocation of funds.

Investment in public securities, divided into three classes, has been permitted in the mutuals for many years. The increase in the federal debt during the 1930's, concomitant with the decline in mortgage demand, has given federal securities a dominant place in the portfolios of savings banks, despite their low yield. Their high marketability facilitates their rapid liquidation in the event of an increased demand for mortgage loans. State and municipal bonds have not grown so rapidly as federals, and have even declined in recent years in the portfolios of some banks. Like the federals, they may serve as an outlet for funds which cannot be more profitably invested, but do not constitute a limitation upon mortgage lending.

Railroad bonds and notes were for many years a favorite bank investment, and during the 'twenties their high yield may have exerted some restriction on mortgage investments, as the large holdings of railroad bonds in certain years indicates (see Table III, above). Street railway bonds have been declining for several years, and it is unlikely that they will ever be an important investment in the future. Since about 1930, neither railroad nor street railway bonds offered any competition with mortgage loans in the allocation of bank funds. Since 1834, savings banks have been permitted to invest in bank stocks. During the Civil War these stocks were especially profitable, but their popularity has declined, and since

1908 they have rarely comprised as much as two per cent of bank assets.

Next to securities, personal loans are the largest potential competitor with mortgages for investment funds. Since 1834, Massachusetts savings banks have been permitted to lend on three-name paper (one principal and two surety promissors). No distinction is made between collateral and indorsed loans, but the sum of these may not exceed one-third of deposits and income. Not more than one per cent of deposits may be lent to a single individual or corporation. Once very popular, personal loans have declined in the mutuals, especially since 1920. They rose temporarily in 1928 and 1929 under the influence of generally high money rates, but have declined sharply since that time, until by 1940 they constituted a negligible proportion of total assets. Personal loans are inferior to mortgage loans: their yield has (with the exception of a few years) been consistently lower, and the ratio of loss has been greater than on mortgage loans. The chief reason for their decline after 1929, however, has been the expense of handling them. After 1928, the number of such loans increased, while their average size decreased, necessitating greater clerical labor. They are now regarded more as a service than as a source of profit. Since about 1930, personal loans have not been a serious competitor of mortgage loans. The dominant investment is now the securities portfolio.[3]

The Massachusetts law contains no provision as to the amount of available cash and cash items to be held by savings banks; it simply limits the amount held on deposit in correspondent banks. The lack of a statutory cash reserve and the ability of mutuals to get along with very little cash (see above), makes available almost the entire amount of assets for income-yielding investments, including mortgage loans.

Finally, the mortgage portfolio is affected by the investment qualities of the mortgage itself. Mortgage loans are high in yield,

3. In trust companies it was otherwise, however. During the early 'twenties personal loans absorbed about half the assets of local trust company savings departments, and overshadowed the mortgage portfolio. But personal loans have declined in trust companies, and although in 1940 they constituted a larger proportion of total assets than was the case in the mutuals, they no longer can be considered a major limitation on mortgage investment. Credit unions invest the bulk of their funds in personal loans to members, and only in these institutions do personal loans substantially restrict the amount of funds available for mortgages — in this case, junior mortgages.

but suffer in liquidity and shiftability. Since the creation of the
Mutual Savings Central Fund in 1932, mutuals in Massachusetts
have been able to discount their mortgage loans and thus maintain
shiftability in the face of adverse real estate conditions. Moreover,
FHA insurance now increases the marketability of mortgage paper,
and amortization facilitates the spacing of mortgage maturities.
During most of the period covered in this study, however, mortgage
insurance was unknown, and amortization was rare. The discount
and borrowing facilities of various government agencies were
instituted only during the past decade. Bonds can still be spaced
as to maturity much more readily than the amortized mortgage,
and the old straight-term, unamortized loan was completely illiquid.
The utter lack of liquidity or shiftability of mortgage investments
during the 'twenties, and the relative weakness of mortgages in
this respect even today, has circumscribed the amount of mortgage
investment. It has undoubtedly influenced the statutory limitation
on mortgage lending, and has also necessitated the investment of
a portion of bank assets in highly marketable securities. (Until
recently it has also kept commercial banks from investing freely
in mortgages.) Were mortgage loans safer, more liquid and more
shiftable, both the banks and the regulating officials would have
been willing to invest a larger proportion of total funds in them,
given an adequate demand. In earlier years, and even today,
mortgages alone did not meet all the requirements of sound invest-
ment policy, and the need for maintaining safety, liquidity,
shiftability and yield led to diversification of investments, with
alternative portfolios sometimes competing for funds against the
mortgage portfolio.

III. MORTGAGE INTEREST RATES

From the annual reports of the bank commissioner an analysis
was made of average rates, modal rates, highest rates and lowest
rates on mortgages held in the five banks annually from 1920
through 1940. These are contract rates and do not show either the
total cost of the loan to the borrower or the net return to the banks,
but they do show the long-term trend. Table VI presents average
rates, weighted according to the amount of mortgage loans held at
each rate in each bank. Chart VI shows the proportion of total
loans held at, above, and below the modal rate.

Average rates exhibit a marked difference in the two decades.

Hovering around 5.94 per cent throughout the 'twenties, the rate dropped abruptly to 5.49 per cent in 1931, after which it continued to fall gradually to 4.86 per cent in 1940. From 1920 through 1930 some 90 per cent of all loans paid 6 per cent. From 1931 through 1940 the most typical rate was 5.5 per cent. The proportion of total loans bearing this rate declined from 80 per cent in 1931 to 49 per cent in 1940, however, indicating the general decline in rates.

TABLE VI

AVERAGE INTEREST RATES ON MORTGAGES HELD IN FIVE SAVINGS BANKS

1920	5.92%	1927	5.94%	1934	5.19%
1921	5.92	1928	5.94	1935	5.08
1922	5.94	1929	5.95	1936	5.02
1923	5.95	1930	5.94	1937	5.02
1924	5.95	1931	5.49	1938	4.98
1925	5.94	1932	5.38	1939	4.91
1926	5.94	1933	5.37	1940	4.86

Source Weighted averages computed from 105 bank statements in Annual Report Part I, 1920–40

Except in 1931, when most rates were reduced, the highest rate was never more than 0.5 per cent above the modal rate, and only a small proportion of mortgages paid interest at more than the modal rate. The lowest rate, on the contrary, has fluctuated considerably in response to changed economic conditions. During the short but severe depression of 1921, a few loans were held at 3 per cent. From 1922 through 1930 the five banks held no loans at less than 5 per cent. In 1931 some loans were held at 3 per cent; in 1932 the lowest rate fell to one per cent; and from 1935 through 1940 some loans were held at 0.25 per cent. There were undoubtedly some loans which paid no interest in those years, but the bank commissioner's reports lack data on this point. The proportion of total mortgages paying interest at less than the modal rate remained about 10 per cent during the 'twenties. Defaults during the depression years and the general reduction of most rates during the 'thirties increased this proportion to 50.1 per cent in 1940. In that year almost half the total loans, by value, were paying 5.5 per cent; almost 28 per cent of the loans were paying 5 per cent; and
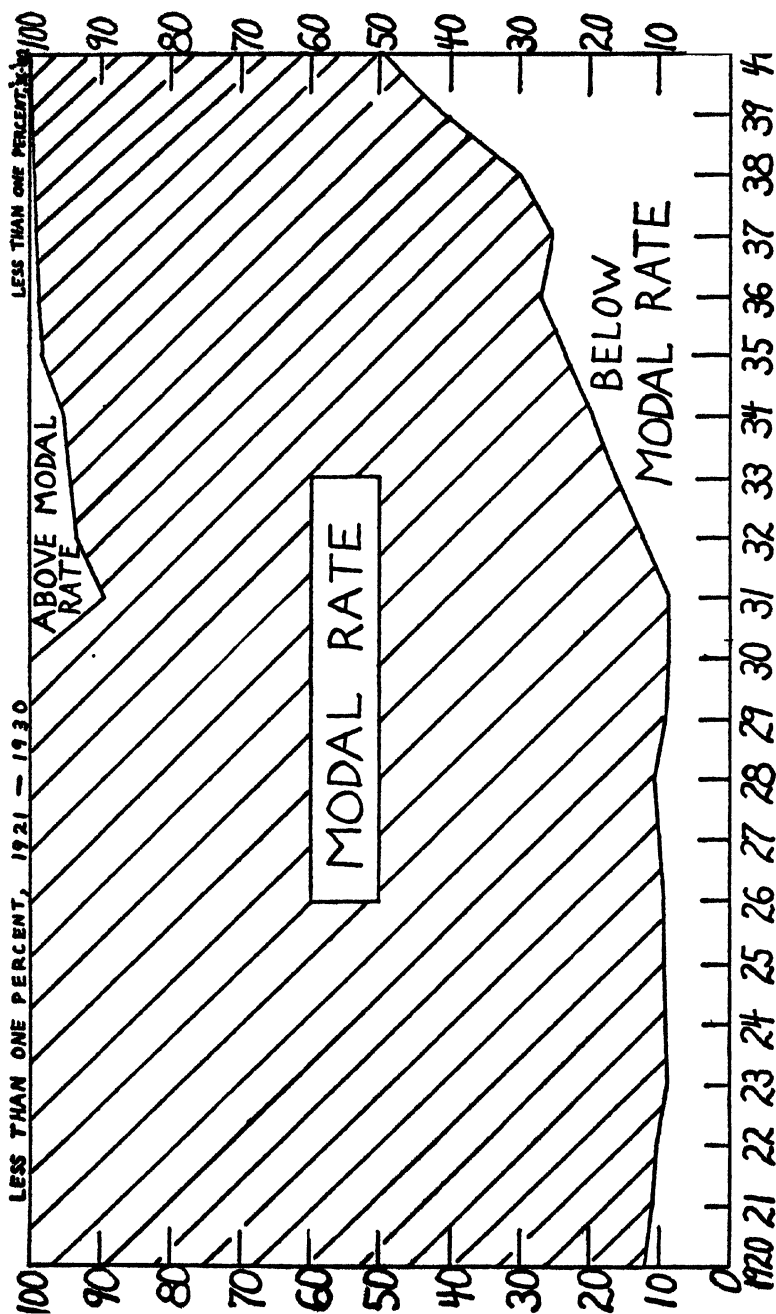
CHART VI

PER CENT OF MORTGAGES HELD AT, ABOVE, AND BELOW MODAL RATE

over 22 per cent of the total were paying less than 5 per cent interest.[4]

The most significant fact revealed by this analysis is the stability or rigidity of savings bank mortgage rates. During the twenty-one year period, the modal rate changed but 0.5 per cent, and the average rate varied only 1.09 per cent. This stability is due, in part, to the long-term nature of the mortgage contract. A loan made at a certain rate will continue to pay that rate years afterward, even though the underlying conditions of supply or demand are changed. Moreover, a greater variation would probably be found in rates on *new* loans made each year. The variation in rates on new lending is partly obliterated in an average rate and especially in a modal rate based on all loans held, since new loans in most years constitute but a small proportion of total loans. Nevertheless, the rates do show remarkable stability over many years. This stability suggests the influence of "custom," fear of "spoiling the market," and other factors of market imperfection, rather than a sensitivity to the forces of supply and demand.[5] When the small variations in rates are compared with the considerable fluctuations in the supply of and demand for mortgage funds as reflected in the changes in volume of mortgage lending (see Section II above, especially Table IV), little relation can be seen between the price of mortgage loans and the underlying market forces.

Another explanation of the stability of rates may be that the supply rather than the price of mortgage funds is the chief variable factor. With a rate more or less stable, an increase in demand may call forth a larger supply of funds from alternative investment

4. An analysis of the 13,612 loans held by the five banks in 1940 revealed that only 0.14 per cent of all mortgages, by value, were paying 6 per cent interest; 49.75 per cent of the loans bore a 5.5 per cent rate; and about half the loans bore a rate of 5 per cent or less. Twenty per cent of all loans bore a rate of 4 per cent or less; 11.4 per cent carried a rate of 3 per cent or less; 3.2 per cent bore a rate of 2 per cent or less; and 1 per cent of all loans, by value, paid a rate of 1 per cent or less. Computed by the writer from data in Annual Report, Part I, 1940. The extremely low rates are "distress rates" on mortgages technically in default but still held in the portfolio. They represent an adjustment on the part of the banks to ease the burden on borrowers in distress, in the hope of tiding them over until they can pay the current market rate of interest. Such an adjustment is often considered temporary, and preferable to foreclosure.

5. See my article, "Imperfect Competition in the Mortgage Market," *Southern Economic Journal*, Vol. X, April 1944, pp. 275–291.

outlets, without requiring an increase in the rate. This is so because mortgage yields have nearly always been considerably more attractive than yields on alternative investments (see Section IV, below). Conversely, when mortgage demand shrinks, funds are channeled into alternative investments (chiefly government securities in recent years), without any significant drop in the mortgage rate. This relative stability of rates in times of depression or of stagnant real estate activity may be due to fear of "spoiling the market" in the future, when demand picks up, or it may be due to realization, on the part of lenders, that in times of depression and slack housing demand the demand for funds is not sufficiently elastic to be stimulated by a drop in financing costs alone. Recovery in the demand for housing, and therefore in the derived demand for mortgage loans, usually awaits other factors in addition to favorable interest rates, such as low construction costs, low materials and labor costs, rising rentals, and reduced vacancy ratios.

In times of distressed economic conditions, moreover, the discriminatory price policies of some lenders permit greater stability of the rate structure than would otherwise be the case. Under a one-price policy, a bank might have to reduce rates on *all* mortgages in order to save some of them from outright default, or it would have to maintain the one rate at its original level and foreclose promptly on all defaulted loans. Under a discriminatory price policy, on the other hand, rates may be considerably reduced on some loans in order to nurse them through a depression without resorting to foreclosure, and at the same time the rates on sound loans may continue with little or no adjustment. This was demonstrated above. If all loans bore the same interest rate, sound loans as well as weak ones would have to be readjusted in order to save a few weak loans from outright default and foreclosure.[6] An alternative policy, serving the same end, is to refinance weak loans at the usual interest rate, but for a longer period of years, so as to reduce the monthly payment.

6. Consider, as an analogy, the effect on the railroad rate structure, if the price of wheat should decline drastically and, in order to maintain this portion of the traffic, rates on all commodities were lowered. With a discriminatory rate policy, such a sacrifice on all business is not necessary, and the variation of rates on part of the business permits greater stability of the rate structure as a whole. (This analogy disregards the influence of fixed costs.)

IV. COMPARISON WITH YIELDS ON OTHER INVESTMENTS

Table VII and Chart VII compare mortgage rates with rates on personal loans and yields on securities owned. In each case the average rate (published in the annual report of the bank commissioner) was weighted by the writer according to the amount of investment of each type in each bank. The average dividend rate is also included for comparison with the yields on loans and investments.

Throughout the two decades the mortgage rate was substantially above the securities yield. It also exceeded the rate on personal loans, except in 1920, 1921 and 1929. The falling yield on personal loans to 1927 probably gave the banks added motive for investing in mortgages during the real estate boom. Improving security yields toward the end of the 'twenties offered a good

TABLE VII

YIELDS ON LOANS AND INVESTMENTS COMPARED WITH DIVIDEND RATE, FIVE SAVINGS BANKS

Year	Mortgages	Personal Loans	Securities	Dividends
1920.........	5.92%	7.24%	4.38%	4.50%
1921.........	5.92	6.68	4.48	4.50
1922.........	5.94	5.10	4.53	4.50
1923.........	5.95	5.28	4.56	4.50
1924.........	5.94	5.00	4.59	4.50
1925.........	5.94	4.75	4.57	4.50
1926.........	5.94	4.88	4.62	4.50
1927.........	5.94	4.75	4.66	4.50
1928.........	5.94	5.71	4.67	4.50
1929.........	5.95	6.54	4.73	4.58
1930.........	5.94	4.62	4.69	5.00
1931.........	5.49	4.45	4.61	4.75
1932.........	5.38	5.24	4.56	4.05
1933.........	5.37	5.02	4.46	3.80
1934.........	5.19	5.14	4.25	3.25
1935.........	5.08	4.97	3.93	3.00
1936.........	5.02	4.85	3.47	2.75
1937.........	5.02	4.57	3.24	2.50
1938.........	4.98	4.58	3.18	2.50
1939.........	4.91	4.47	3.07	2.50
1940.........	4.86	4.44	3.02	2.36

Source: Weighted averages computed from 105 bank statements in Annual Report, Part I, 1920–40.
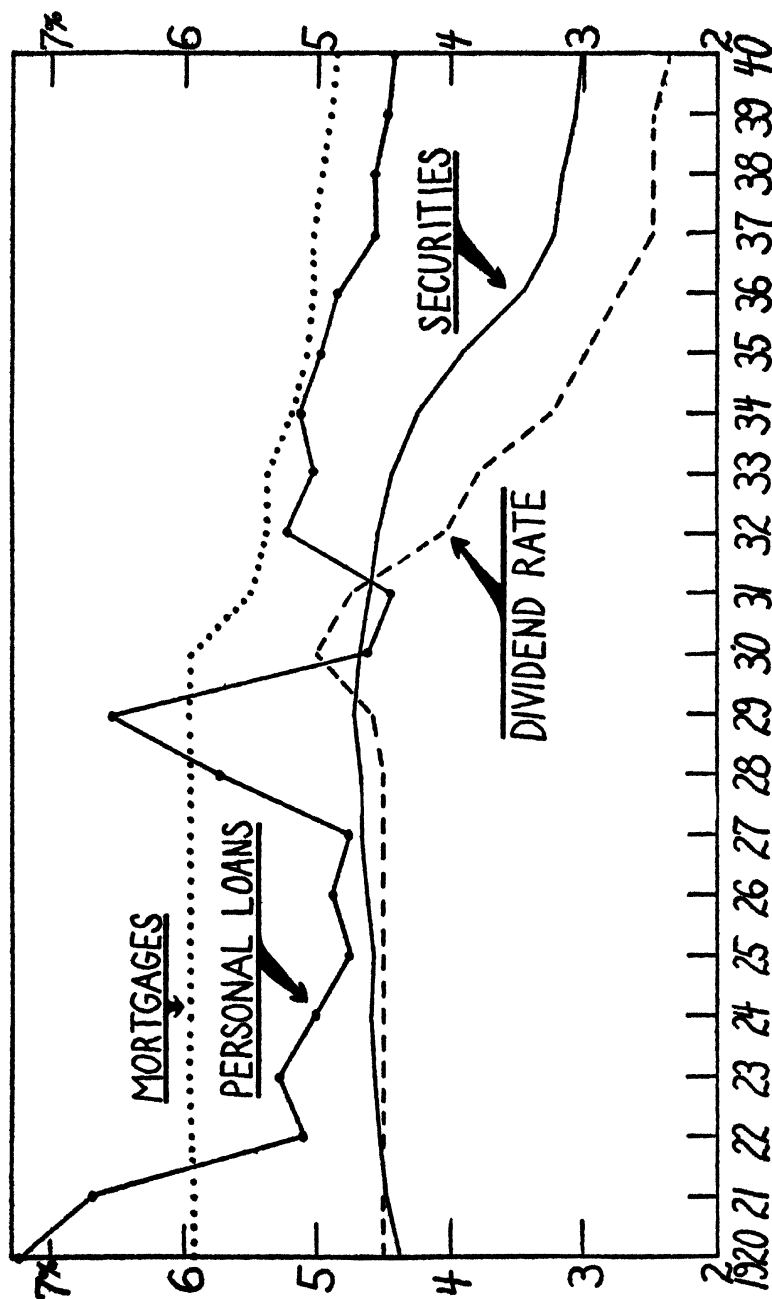
CHART VII

Yields on Loans and Investments and Dividend Rate

investment outlet to banks which had reached the statutory limit of mortgage investment. As the demand for mortgage loans declined during the 'thirties, a growing proportion of funds was placed in government securities. Chart VII reveals that, although security yields have fallen, since 1931 the spread between the dividend rate and the securities yield has been growing. Thus it was profitable for these banks to invest heavily in securities during the 1930's, although not so profitable as mortgage loans. Finally, although personal loans bear a higher interest rate than securities, they have declined in importance as their cost of handling has increased. Since 1934 the difference between the rate on mortgages and personal loans has been growing.

In 1940, the yield on mortgage loans in the five Worcester mutuals exceeded the dividend rate by 2.50 per cent; the yield on personal loans was 2.08 per cent above the dividend rate; and securities yielded 0.66 per cent above the dividend rate. This last spread, though small, is probably more than sufficient to absorb the costs connected with the securities portfolio. The ratio of total expenses plus taxes to total assets of the five Worcester mutuals was only 0.44 per cent in 1940.[7] The securities portfolio, being partially tax exempt and involving little servicing expense, probably accounts for the smallest portion of overhead costs. Little risk is involved, since the portfolio is so heavily loaded with government securities.

If a bank's fund-conversion policy were based solely on relative yield, and if there were no legal restrictions on investment, and if, furthermore, the demand for a bank's funds in each direction were unlimited, the greatest proportion of funds would be invested in mortgage loans, the second largest amount in personal loans, and the smallest amount (consistent with safety and shiftability) in securities. But these are heroic assumptions. Legal restrictions limit the mortgage portfolio to 70 per cent of deposits, and require some investment in securities to insure safety and shiftability. The main reason why the investment pattern in 1940 did not conform to relative yields was that the demand for mortgage loans *at then current rates* was not sufficient to absorb the funds which would

7. Weighted average of expenses (including salaries, rent, advertising and other expenses of management) plus State taxes, computed by the writer for the five Worcester banks for the fiscal year ending October 31, 1940. This compares favorably with the ratio for all Massachusetts mutuals, which averaged 0.46 per cent of assets. Annual Report, 1940, Part I, p. xvi.

normally be available at the attractive spread of 2.50 per cent. Personal loans being expensive to handle and the demand for them being small, the final result was that the largest proportion of these banks' assets was invested in the lowest-yielding portfolio — that of securities. Should the demand for mortgage loans in Worcester increase substantially in the years ahead, and should the mutuals receive a fair share of that demand, the yield differential between mortgage loans and all other loans and investments (even at lower mortgage rates than those currently prevailing) offers a continuing incentive to direct more and more funds into the mortgage portfolio.

V. Summary and Conclusions

Savings in Worcester's five mutual savings banks fluctuated with the business cycle, but enjoyed a secular growth since 1920. Mortgage loans absorbed a growing proportion of these savings during the early and middle 'twenties. The 1930's witnessed a great expansion in bank holdings of federal bonds and a decline in mortgages and personal loans. The volume of mortgage loans held fluctuated with the general business cycle, while the volume of new loans made coincided with the real estate cycle. Legal and economic factors have influenced mortgage loans and alternative investments. The legal mortgage-deposit ratio remains at 70 per cent, but debt-value ratios have risen, permitting banks to lend more on each loan. The broadening of the legal list of securities after 1926 did not restrict mortgage investments; it simply absorbed funds which could no longer be invested in mortgages after the peak of the real estate cycle had been passed. The rising yield of public utility bonds in the late 'twenties probably would have made them compete further with mortgages, had the demand for the latter not fallen off before the general decline in business and industrial activity. The swollen portfolio of government bonds in the 'thirties was not a cause but a *consequence* of the shrinkage in the mortgage portfolio. The falling yield of utility, railroad and government bonds during the 'thirties enhanced the relative attractiveness of mortgage investments, and probably would have stimulated the latter, had the demand been adequate or had lenders been aware of the elasticity of demand for mortgage loans.[8]

8. I have seen no adequate analysis of the elasticity of demand for mortgage loans. It is probable that this elasticity is greater than most lenders assume, and that a reduction in financing costs would have a more than

The liberal statutory provision allowing investment up to one-third of deposits in personal loans might have restricted mortgage investment, had the demand for personal loans been sufficiently great and the yield sufficiently high to compete with mortgages. This was not the case, however. In sum, the potential supply of mortgage money was increased in the 1930's by the growth of savings and the declining attractiveness of alternative investments, but this potential supply was not so fully utilized as it was during the real estate boom of the 'twenties.

Although mortgages provide an adequate and stable yield, they were (and to a lesser extent still are) deficient in liquidity and shiftability. This deficiency has been a dominant motive behind statutory restrictions on mortgage investments, and it necessitates some investment in more shiftable assets such as bonds, thereby indirectly limiting the mortgage portfolio. These deficiencies have been remedied somewhat in recent years by amortization and insurance and by the creation of facilities for rediscounting and borrowing upon mortgage paper. As the secondary market for such instruments becomes wider, we may expect an increase in the supply of institutional funds available for home financing. In regions where they are prominently established, the mutual savings banks are in an excellent position to take the lead in this respect.

Mortgage rates were stable during the 'twenties but declined gradually during the 'thirties. This decline is certainly not impressive, in view of the decline of interest rates generally, the broader sources of available funds, the growing safety and marketability of insured, amortized mortgages, and the efforts of government agencies to create a sounder market. Nevertheless, mortgage rates charged by the five mutual savings banks were below those of competing lenders throughout the period studied. Although all loan rates have fallen since about 1930, the dividend rate has fallen still more sharply. As a result, the spread between the cost of attracting savings and the gross yield on all investments has grown. The largest gross yield is returned by mortgages, followed in turn by personal loans and securities. However, the demand for mortgage funds declined during the 'thirties and personal loans

proportionate effect in enlarging the demand for home ownership and residential construction and the derived demand for mortgage funds. Opinions were presented to that effect in the Hearings before the Temporary National Economic Committee (Vol. 11, Construction Industry), but a careful inductive study is needed before we can proceed confidently in this bewilderingly complex field. Cf. note 9, below.

have become expensive to handle; consequently the securities portfolio ranks first in size. Thus the allocation of invested funds does not conform to relative yields, even after allowance for the factors of safety and liquidity. The yield differential of mortgage loans over securities offers the banks a continuing incentive to expand the former and contract the latter.

Despite the decline in mortgage rates during the 1930's, the demand for mortgage money has not kept pace with the available supply. A further reduction of rates below those currently prevailing would probably stimulate the demand for home ownership and for mortgage loans.[9] This might well result in maximizing the net revenue of the savings banks, which would do a volume business at smaller unit earnings. To some extent the banks before the war attempted to maintain mortgage rates in order to reimburse themselves for the decline of revenue on other investments. In view of the general decline of money rates,[1] and the avowed objective of many lenders to stimulate home ownership, the construction industry and the utilization of idle funds, such a policy cannot be justified now. The undesirable alternative to the full utilization of funds in mortgage loans is a swollen portfolio of government bonds.

9. Of course, now and for some years to come the problem is lack of building materials and labor, not lack of funds. As long as shortages persist, even a substantial reduction of mortgage interest rates would not greatly increase house construction. It would simply increase transfers of existing houses, at inflated prices. But the conclusion above may still be relevant in a future depression, or following an era of overbuilding, or in connection with the problem of providing homes for families in modest-income brackets. On the other hand, as mortgage rates decline, financing costs will become less important in the total costs of home ownership, and the elasticity of demand for loans will be diminished. And with every progressive reduction in mortgage rates, the other factors in the cost of building construction and home ownership become relatively more important.

1. This decline has continued since 1940. According to a recent survey, the average yield on high-grade corporate bonds (Moody's AAA) declined from 2.84 per cent in 1940 to 2.48 per cent in 1946; the average interest earned by leading life insurance companies has fallen from 4.11 per cent in 1940 to an estimated 3.65 per cent for 1946; and the average interest ("dividends") paid by mutual savings banks to depositors has fallen from 2.26 per cent in 1940 to 1.80 per cent in 1946. During the same period there was a great expansion in savings of all types. Time deposits of 531 mutual savings banks in seventeen states increased from less than $10.5 billion in 1940 to over $15.3 billion in 1946. New York Times, March 24, 1946, p. F 1.

PAUL M. GREGORY

UNIVERSITY OF NORTH CAROLINA
GREENSBORO

FORWARD SHIFTING OF A PAYROLL TAX UNDER MONOPOLISTIC COMPETITION[1]

SUMMARY

I. The incentive argument for merit rating, 267. — Definitions, 268. — II. Monopoly profit before and after the imposition of contribution costs, 269. — Implications for forward shifting of contribution costs, 270. — III. Adequacy of profit reductions, 274. — Effect of different contribution rates upon a given monopolist, 274 — Effects of elasticities of demand and slopes of marginal cost functions, 276. — Conclusions, 283.

I

The term merit rating or experience rating has been employed in the field of unemployment compensation to indicate differential rates of contribution by employers (payroll tax) to the unemployment compensation fund in a given state. In some states, for example, it is possible for some employers to pay contributions at one rate and other employers to pay contributions at a higher or lower rate. The difference in rates depends upon the standard of performance of each employer or each group of employers with respect to employment stabilization. What stabilization of employment means or should mean, or how it should be measured, and to what extent employers acting individually can achieve stabilization, are controversial issues which will not be discussed in this article.

Proponents of experience rating argue in support of the system that it will act as an incentive to employers to stabilize their employment. It will do this, they hold, because the achievement of stabilization will result in lower contribution costs.[2] On the other hand, if an employer fails to stabilize his employment, the resulting increase in contribution costs will place him at a competitive disadvantage. Therefore, in order to earn more or lose

2. See Bryce M. Stewart, Planning and Administration of Unemployment Compensation (New York: Industrial Relations Counsellors Inc., 1938), pp. 492–497. Herman Feldman and Donald M. Smith, The Case for Experience Rating in Unemployment Compensation and a Proposed Method (New York: Industrial Relations Counsellors Inc., 1939), pp. 3–4. Richard A. Lester and Charles V. Kidd, The Case Against Experience Rating in Unemployment Compensation (New York: Industrial Relations Counsellors Inc., 1939), p. 5.

less, he would try to introduce measures aiming at stabilization.

The validity of this argument rests, in large part, upon the premise that the penalty of a higher contribution rate (payroll tax) will change the cost of production relative to receipts to such a degree that profit will be decreased sufficiently to induce the employer to attempt to improve his rating. This premise assumes, among other things, that the burden of the contribution costs cannot be shifted in total or in large enough degree by the employer. It is not clear, however, that the burden cannot be shifted forward by increasing the price of the finished product, or backward by decreasing the prices which he pays for the factors of production. The purpose of this study is to examine the possibilities of shifting the burden forward.[3]

As in most problems, the analysis of the incidence of the contribution costs first necessitates rigorous definitions of certain concepts. Just when is the burden of the contribution costs "shifted"? The following definition will be used in this article. The burden of the contribution costs not only refers to the dollar amount of the cost but also to the "effect" or "pressure" of the cost. The burden is measured by the decrease in the net income of the person on whom the tax falls in the first place. *Thus, complete shifting occurs only when the net income (money) of the person upon whom the tax is placed originally is not decreased as a result of the tax.*[4]

This definition of shifting the burden seems adequate for the analysis of the problem of the possibilities of shifting contribution costs forward by raising prices. The problem before the producer is not simply to increase average price to cover the contribution cost per unit produced, regardless of the number of units sold, but rather to change total receipts and total costs in such a manner as to maintain his net income. The real query is: can the producer maintain the same total net income *after* the imposition of contribution costs as he had *before* the imposition of contribution costs?

3. Some factors are in operation which tend to decrease the significance of backward shifting, particularly with regard to shifting the burden back to wage earners through lower wages. The importance of backward shifting to labor is probably decreased by the increase in the bargaining power of organized labor. In the case of unorganized labor, minimum wage legislation, old age benefits, and even relief, probably tend to lessen the possibility of backward shifting.

4. This definition is in accord with the definition used by Bastable in his discussion of shifting the burden of a tax. C. F. Bastable, Public Finance (3d ed., New York, 1927), p. 361.

In the process of forward shifting, which is dealt with below, the burden of contribution costs is transferred by increasing the selling price of the product. (Labor might suffer by a decrease in employment but, implicit in the problem posited, not by a decrease in wage rates). The degree to which the burden can be shifted through such an increase in prices depends upon the elasticity of the demand for the commodity and the slope of its marginal cost curve.[5] In addition, it depends upon whether or not the employer is selling the commodity under conditions of monopolistic or perfect competition. In this paper the problem of forward shifting under monopolistic competition will be discussed. A somewhat general solution for the individual firm will be attempted, and some special cases assuming varying elasticities of demand and varying cost conditions for individual firms will be analyzed.

II .

The price of a particular commodity produced under monopolistic competition is a function $f(q)$ of the quantity offered for sale q.[6] The function $f(q)$, the demand curve, is assumed to be a negatively-sloped curve. The total wage cost of production is represented by the function $\varphi(q)$, which is a positively-sloped curve; the total material cost is represented by mq, where m is a constant designating the material cost for one unit of output; and the total overhead cost is represented by a constant k. The total profit[7] or net income $P(q)$ of a monopolist can be stated in the form

(1) $P(q) = qf(q) - \varphi(q) - mq - k;$

and the maximum profit would obtain where

(2) $P'(q) = qf'(q) + f(q) - \varphi'(q) - m = 0,$

in which $P'(q)$, $f'(q)$ and $\varphi'(q)$ denote the first derivatives of $P(q)$, $f(q)$ and $\varphi(q)$, respectively. The equation for the total profit of the

5. Cf. Edison L. Bowers, "Social Security Program," *American Economic Review*, XXVIII, No. 1, Supplement (December, 1937), p. 140.
6. Throughout this analysis the quantity offered for sale and the quantity produced are assumed to be the same and are represented by q.
7. This article deals with the income function under a given set of conditions for one period. For a discussion of maximum present value of income after tax for more than one accounting period, see Philip D. Bradley, "The Direct Effects of a Corporate Income Tax," this JOURNAL, Vol. LVI, (1942), p. 639.

monopolist after the imposition of the unemployment compensation contribution rate r, which is applied to wages, becomes[8]

(3) $P_1(q) = qf(q) - \varphi(q) - mq - k - r\varphi(q)$

Subtracting equation (1) from (3) we have $P_1(q) - P(q) = -r\varphi(q)$ which equals

(4) $P_1(q) = P(q) - r\varphi(q)$.

Before the increased cost due to contribution costs, maximum profits obtained where $P'(q) = 0$, but after the increased cost due to contributions, maximum profits would obtain where

(5) $P_1'(q) = 0 = P'(q) - r\varphi'(q)$ or $P'(q) = r\varphi'(q)$.

That is, after the imposition of the contribution costs, the monopolist would produce up to the point where his marginal profit $P'(q)$ before the imposition of the contribution costs equals the contribution rate r times his marginal labor cost, which, to change the wording, is the same as the amount of the contribution cost on the last unit produced.

Assume that the net income or profit function could be represented by a second degree curve which increases and then decreases, that is, the profit function would be positively sloped up to the point of optimum (relative to profits) output and then would become negatively sloped. The marginal profit function, therefore, will be negatively sloped and have positive values as output increases up to the optimum output, after which point it will have negative values. Under these conditions, the profit and marginal profit functions, before and after contribution costs, can be represented graphically as in Figure I.

Before the increased cost, the monopolist made his maximum profit at an output of q_o, which is the root of the equation $P'(q) = 0$. After contribution payments are made, however, he maximizes profits at an output of q_{o-u}, which is the root of equation

$$P_1'(q) = 0 = P'(q) - r\varphi'(q).$$

From equation (4) it follows that for any given quantity sold, q, the net income or profit before the pay roll tax is greater than that afterwards. Hence the amount of maximum profit is greater

8. This is short-run analysis in which it is assumed that the demand function and the material costs, wage rates, and overhead costs, in short, all the various prices of the factors of production, do not change as a result of the imposition of the unemployment compensation contribution or for any other reason.
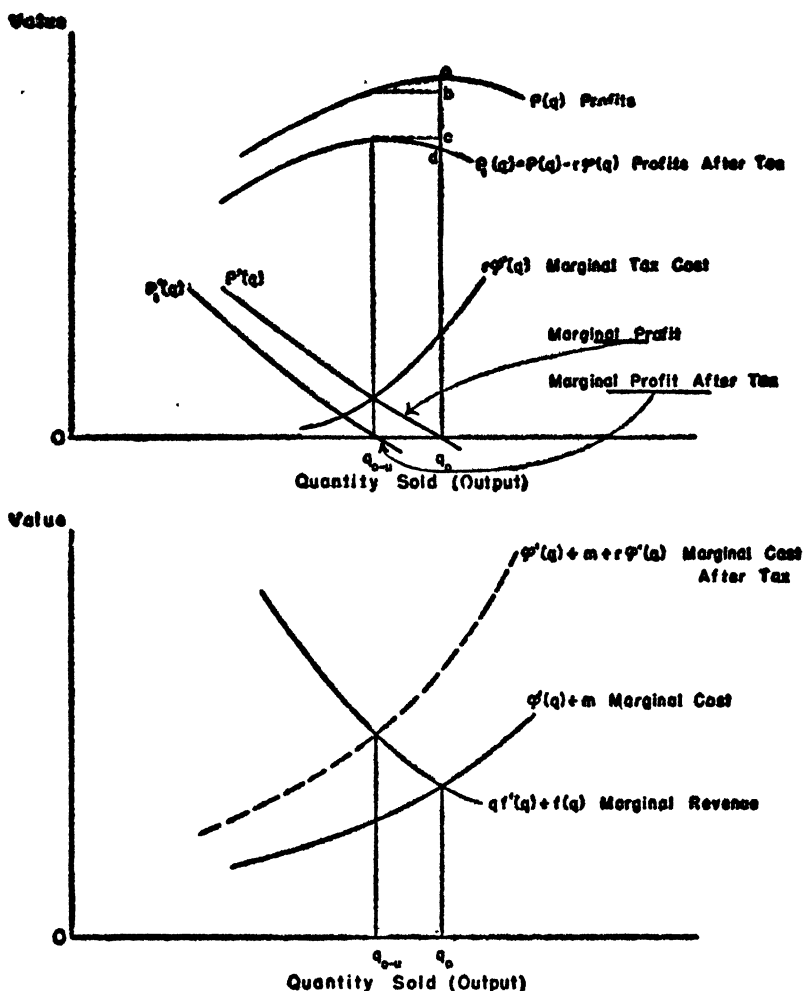
FIGURE I

PROFIT AND MARGINAL PROFIT
BEFORE AND AFTER CONTRIBUTION COSTS (TAX)

before than after contribution payments. As a result, in the case of monopolistic competition, the burden of the contribution costs cannot be shifted completely forward. Moreover, it follows from equations (2) and (5) that the quantity produced and sold which yields maximum profits after contribution payments will always be

less than before[9] (see Figure I), except for the theoretical case where wage costs do not vary with output. In addition, since the demand function is negatively sloped, the price of the commodity will be higher at the smaller output.[1]

The difference in the net income to the producer is

$$(6) \qquad \int_o^{q_0} P'(q) - \left[\int_o^{q_0-u} P'(q) - r \int_o^{q_0-u} \varphi'(q) \right]$$

which is always a positive quantity.

Since $\int_o^{q_0-u} P'(q) < \int_o^{q_0} P'(q)$, the decrease in net income is greater than $r \int_o^{q_0-u} \varphi'(q)$. Also, $r \int_o^{q_0-u} \varphi'(q)$ is equal to the amount

9. Cf. Seymour E. Harris, Economics of Social Security (New York, 1941), p. 381.

1. The increase in price after contribution payments are made can be demonstrated using the method suggested by A. Cournot, The Mathematical Principles of the Theory of Wealth (New York, 1929), pp. 61-63. Let p equal price, $f(p)$, quantity sold; $\varphi(q)$, total variable cost (consists of wages only); and k, overhead cost. Assume, furthermore, that variable cost consists only of wages.

(a) $pf(p) - \varphi(q) - k = $ profits.

(b) $f(p) + f'(p)[p - \varphi'(q)] = 0$. The root of this equation gives the value for p which renders profits a maximum. With the imposition of the unemployment compensation contribution rate r, equation (b) becomes:

(c) $f(p) + f'(p)[p - \varphi'(q)(1+r)] = 0$.

Assuming that p_0 is the root of equation (b), that $\varphi'(q)$ can be expressed as a function of p, say $\psi(p)$, and that δ represents the amount of the change in price occasioned by the new marginal cost which resulted from the pay roll tax, equation (c) can be written as follows:

$$f(p_0+\delta) + f'(p_0+\delta)[(p_0+\delta) - \psi(p_0+\delta)(1+r)] = 0.$$

Expanding by Taylor's theorem, neglecting higher powers of δ and r, and subtracting equation (b) after substituting $\psi(p_0)$ for $\varphi'(q)$ in equation (b), the result is:

$$\frac{\delta}{r\psi(p_0)} \qquad \frac{f'(p_0)}{f'(p_0)[2 - \psi'(p_0)] + f''(p_0)[p_0 - \psi(p_0)]}$$

Both numerator and denominator in the second member of the equation above are negative. The numerator is negative because the demand function is negatively sloped, and the denominator must be negative if p_0 renders equation (a) a maximum, since the denominator is the first derivative of equation (b). Therefore δ (amount of the change in the price) must be of the same sign as $r\psi(p_0)$ (contribution payment which would be made on the last unit produced if production did not decrease) which is positive. The increase in the price δ can be expressed as:

$$\delta = \frac{r\psi(p_0)f'(p_0)}{f'(p_0)[2 - \psi'(p_0)] + f''(p_0)[p_0 - \psi(p_0)]}.$$

of money paid as unemployment compensation contributions. Consequently, the decrease in net income to the producer is greater than the amount of his contribution payment for unemployment compensation. It must be kept in mind, however, that the decrease in net income is still less than it would be if the producer did not raise price and decrease the quantity sold, that is, if he continued to sell the quantity q_o at the original price.[2]

The distance *ad* represents the decrease in net income (original burden) if no attempt were made to shift the burden forward (Figure I). The distance *ac* represents the decrease in net income (final burden) after forward shifting. The distance *cd* represents the amount of the original burden shifted forward. Also, the distance *ad* represents the amount of revenue the Government would have received if no attempt were made to shift the burden forward. The distance *bc* represents the amount of revenue received by the Government (contribution costs) after forward shifting. The distance *ab* and *cd* represent the decrease in revenue to the Government after forward shifting. Finally, the distance *ab* represents that part of the decrease in net income after forward shifting which is over and above the amount paid in contribution costs.

2. The decrease in net income equals $\int_{q_{o-u}}^{q_o} P'(q) + r\int_{}^{q_{o-u}} \varphi'(q)$ which

is less than $r\int_{}^{q_o} \varphi'(q)$, the decrease in net income if the monopolist does not increase his price and decrease the quantity he produced. This obtains because if the quantity q_{o-u} renders profits a maximum after contributions have been made, $P'(q)$ must be less than $r\varphi'(q)$ for any quantity greater than q_{o-u} and hence $r\int_{o-u}^{q_o} \varphi'(q) > \int_{q_{o-u}}^{q_o} P'(q)$. If, however, a company produces without any wage cost or with a constant wage cost, namely, where wage cost does not vary with the amount of business, the situation which obtains is somewhat unique. From equations (4) and (5), it is apparent that the quantity sold and the price would not change as a result of the contribution payment. In the case of no wage cost, the company would pay no contribution, and hence its net income would not be changed. In the case of company with a constant wage bill, the quantity q_{o-u} would equal q_o and hence the loss in net income would equal $r\varphi(q)$ where $\varphi(q)$ equals the constant wage cost or the loss in net income equals the full amount of the contribution payments. Price and output would remain the same as before the increased cost. No forward shifting would occur.

III

The critical question whether or not profits would be decreased sufficiently to induce the monopolist to attempt measures by which he would decrease the amount of his contribution costs cannot be answered in general. It is possible to say, however, that all monopolists will have their profits decreased as a result of contribution costs, and that some companies will have their profits decreased relatively more than others. If company A and company B have the same profit function $P(q)$, but company A has a higher marginal cost (assume all variable costs to be wages) and a higher demand function, the net income after the imposition of the contribution costs for company A will be absolutely and relatively less than that for company B. Net income in company A will be absolutely less because $r\varphi_A(q) > r\varphi_B(q)$

for any given quantity q, and as a result

$P_B(q) = P(q) - r\varphi_B(q) > P_A(q) = P(q) - r\varphi_A(q)$, where $P_B(q)$ and $P_A(q)$ represent profits of company B and company A after the imposition of the contribution costs.

Since the above is true for any given quantity q after contributions are paid, the maximum profits for company B will be greater than the maximum profits for company A. The relative decrease in the profits of company B is less, also, than that for company A because both maximum profits were the same before the imposition of the contribution cost. The marginal profit functions before and after the contribution payments for both company A and B can be represented graphically as in Figure II.

Before the unemployment compensation contributions were paid, both companies produced and sold q_o; the net income for both companies was the same, and was equal to the value of the integral of $P'(q)$, where q equals q_o. With the increased cost due to the contribution payments, company A cut quantity sold to q_A, and company B cut to q_B. The net income now for company B, which is the value of the integral of $P_B'(q)$ where q equals q_B, is greater than the net income for company A, the value of the integral of $P_A'(q)$ where q equals q_A.

The above analysis could be used to demonstrate the effects of different contribution rates upon the net income of a given monopolist. Assume that instead of two companies as above, one company is to be analyzed to discover the effect of smaller and larger

$P'_A(q) = P'(q) - r \mathcal{G}'_A(q)$ Marginal Profit Co. A. After Contributions

$P'_B(q) = P'(q) - r \mathcal{G}'_B(q)$ Marginal Profit Co. B. After Contributions

$P'(q)$ Marginal Profit Co. A. And B. Before Contributions

$r \mathcal{G}'_A(q)$ Marginal Tax Cos'
Co. A.

$r \mathcal{G}'_B(q)$ Marginal Tax Cos'
Co. B.

Value

O

Quantity Sold (Output)

q_o q_B q_A
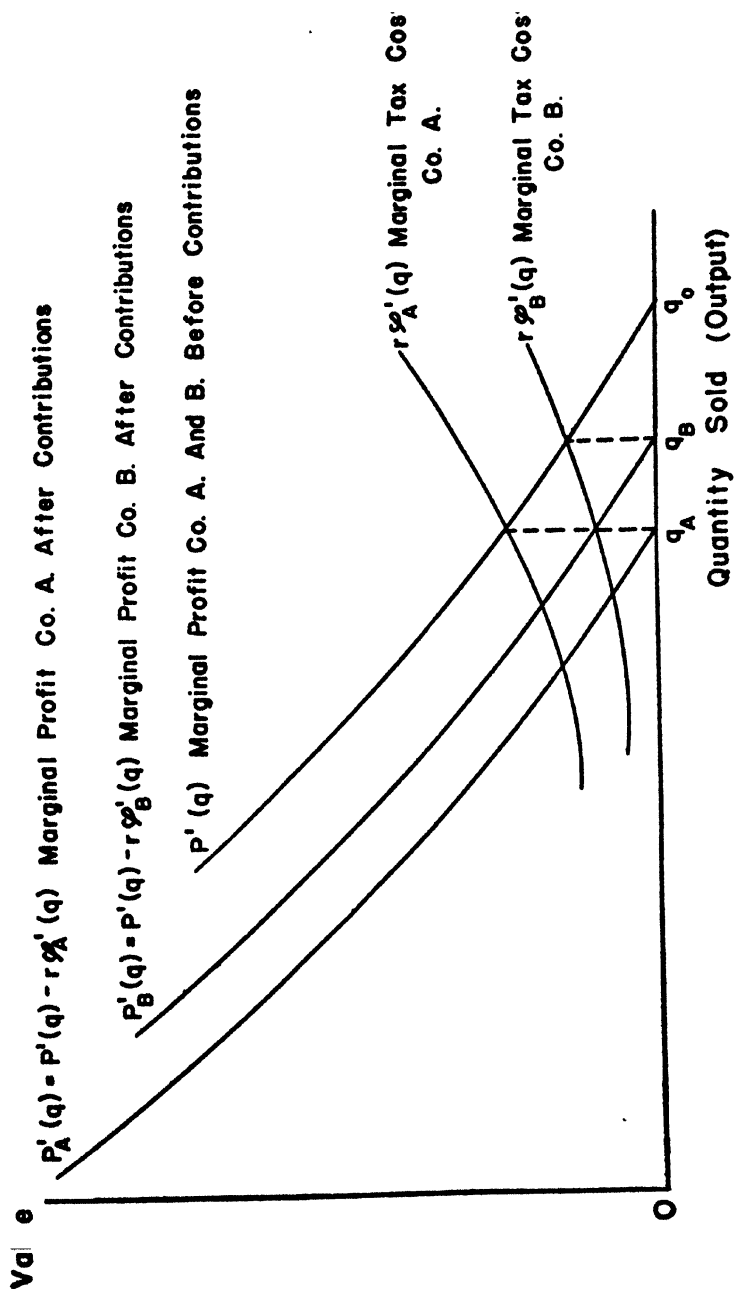
FIGURE II

MARGINAL PROFIT BEFORE AND AFTER CONTRIBUTION COSTS (TAX) IN COMPANY A AND B

contribution rates upon his profit. Designating the larger rate by r_A and the smaller rate by r_B, the following equations represent the two different profit functions for the one company:

$$P_A(q) = P(q) - r\varphi(q), \text{ and } P_B(q) = P(q) - r_B\varphi(q).$$

Since $P(q) = P(q)$ and $r_A\varphi(q) > r_B\varphi(q)$ for any given quantity q, $P_B(q) > P_A(q)$ for any given quantity and, as a result, the maximum $P_B(q) >$ the maximum $P_A(q)$. This means that if a monopolist can regulate his production or employment so that he is assessed at a smaller rate, his net income or profit will be reduced less, quantity will be restricted less, and price will be increased less.

To show the relative effects of various elasticities of demand and various slopes of marginal cost curves on the relative decrease in profits, some specific cases will be taken up.[3] To simplify the analysis, assume constant elasticity of demand. The demand is

$$f(q) = \frac{a}{q^{1/\epsilon}}$$

where a is some constant, q represents quantity sold, and ϵ equals the elasticity of demand, the greater ϵ, the greater the elasticity of demand.

In addition, assume a constant slope of the marginal cost, all variable costs to be wages, and production to be in the stage of diminishing marginal returns. The marginal cost function is

$$\varphi'(q) = \frac{q^\sigma}{b}$$

where b is some constant, q represents quantity supplied, and σ determines in large part, the slope; the greater σ, the greater the slope of the marginal cost curve or the faster costs increase with output. In order to solve for the quantity which renders profits a maximum, it is necessary to substitute the above functions in equation (2). Equation (2) then becomes

3. In analyzing the incidence of taxation, the use of the elasticity of demand is more fruitful than the use of the slope of the demand curve, but the reverse of this is true in the case of marginal cost curves. For discussion of elasticity formula for demand functions, see Alfred Marshall, Principles of Economics (8th ed., London, 1920), Math. App. Note III, p. 839; for supply functions, see Henry J. Bittermann, "Elasticity of Supply," American Economic Review, XXIV (1934), p. 417.

$$(7) \qquad q\left[\frac{-a}{eq^{\left(1+\frac{1}{e}\right)}}\right] + \frac{a}{q^{\frac{1}{e}}} - \frac{q^{\sigma}}{b} = 0.$$

Solving for q, it becomes

$$-\frac{a}{eq^{\frac{1}{e}}} + \frac{a}{q^{\frac{1}{e}}} = \frac{q^{\sigma}}{b}$$

$$\frac{a}{q^{\frac{1}{e}}}\left[1 - \frac{1}{e}\right] = \frac{q^{\sigma}}{b}$$

$$ab\left(1 - \frac{1}{e}\right) = q^{\left(\sigma + \frac{1}{e}\right)}$$

$$(8) \qquad q \text{ or } q_o = \left[ab\left(1 - \frac{1}{e}\right)\right]^{\left(\frac{1}{\sigma + \frac{1}{e}}\right)}$$

The general equation of profits is

$$(9) \qquad qf(q) - \varphi(q) - k = \text{ profits,}$$

and, after substituting the above type of demand and marginal cost function (marginal cost integrated to total variable cost), it becomes

$$(10) \qquad q\left(\frac{a}{q^{\frac{1}{e}}}\right) - \frac{q^{(\sigma+1)}}{(\sigma+1)b} - k = \text{ profits.}$$

Furthermore, substituting the value for q_o (equation 8), which renders profits a maximum, equation (10) becomes

$$a\left[ab\left(1 - \frac{1}{e}\right)\right]^{\left(\frac{1 - \frac{1}{e}}{\sigma + \frac{1}{e}}\right)} - \frac{\left[ab\left(1 - \frac{1}{e}\right)\right]^{\left(\frac{\sigma+1}{\sigma+\frac{1}{e}}\right)}}{(\sigma+1)b} - k = \begin{matrix}\text{maximum}\\\text{profits.}\end{matrix}$$

In the case which has been here assumed, the total variable cost (labor cost) at the point of maximum profits is $\left[\dfrac{1 - \frac{1}{e}}{\sigma + 1}\right]$ times

the total receipts.[4]

After the increased cost due to the unemployment compensation contributions, equation (2) becomes

$$qf'(q)+f(q)-\varphi'(q)(1+r)=0$$

and after substituting the demand and marginal cost functions in question, it becomes

$$q\left[\frac{-a}{\epsilon q^{\left(1+\frac{1}{\epsilon}\right)}}\right] + \frac{a}{q^{\frac{1}{\epsilon}}} - \frac{q^{\sigma}}{b}(1+r)=0.$$

Solving for q it becomes

$$q \text{ or } q_{0\text{-}u}=\left[\frac{ab\left(1-\frac{1}{\epsilon}\right)}{(1+r)}\right]^{\left(\frac{1}{\sigma+\frac{1}{\epsilon}}\right)}$$

The general equation for profits after contribution payments are made is

$$qf(q)-\varphi(q)(1+r)-k=\text{ profits,}$$

and after substituting the above type of demand and marginal cost functions as well as the above value of $q_{o\text{-}u}$, maximum profits become

4. This can be demonstrated in the following manner:

$$\text{Total receipts}=a\left[ab\left(1-\frac{1}{\epsilon}\right)\right]^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}$$

$$\text{Total variable cost}=\frac{\left[ab\left(1-\frac{1}{\epsilon}\right)\right]^{\left(\frac{\sigma+1}{\sigma+\frac{1}{\epsilon}}\right)}}{b(\sigma+1)}$$

$$\frac{ab\left(1-\frac{1}{\epsilon}\right)}{ab(\sigma+1)} a\left[ab\left(1-\frac{1}{\epsilon}\right)\right]^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}$$

$$=\frac{a\left[ab\left(1-\frac{1}{\epsilon}\right)\right]\left[1+\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)\right]}{ab(\sigma+1)}$$

$$=\frac{\left[ab\left(1-\frac{1}{\epsilon}\right)\right]^{\left(\frac{\sigma+1}{\sigma+\frac{1}{\epsilon}}\right)}}{b(\sigma+1)}=\text{ variable cost.}$$

$$a\left[\frac{ab\left(1-\frac{1}{\epsilon}\right)}{(1+r)}\right]^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)} - \left(\frac{1}{b(\sigma+1)}\right)\left[\frac{ab\left(1-\frac{1}{\epsilon}\right)}{(1+r)}\right]^{\left(\frac{\sigma+1}{\sigma+\frac{1}{\epsilon}}\right)} \times$$

$$(1+r)-k = \text{maximum profits.}$$

Under the given conditions, total receipts at the point of maximum profits after contributions are equal to total receipts before contributions divided by $(1+r)^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}$. The same obtains for total variable costs.

To summarize, at the point of maximum profits, the following relationships exist:

(a) Total variable costs before contributions $= \left(\frac{1-\frac{1}{\epsilon}}{\sigma+1}\right)$ times total receipts before contributions;

(b) Total receipts after contributions $=$ total receipts before contributions divided by $(1+r)^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}$;

(c) Total variable costs (including the contributions) after contributions $=$ total variable costs before contributions divided by $(1+r)^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}$.

From the above relationships, it follows that the equation which yields the relative decrease in profit can be written as:

$$(11) \quad \frac{100}{P}\left[\left[R-\left(\frac{1-\frac{1}{\epsilon}}{\sigma+1}\right)R-k\right] - \left[\frac{R-\left(\frac{1-\frac{1}{\epsilon}}{\sigma+1}\right)R}{(1+r)^{\left(\frac{1-\frac{1}{\epsilon}}{\sigma+\frac{1}{\epsilon}}\right)}}-k\right]\right]$$

where $R=$ total receipts at point of maximum profits before contributions; $\frac{\left(1-\frac{1}{\epsilon}\right)}{(\sigma+1)}R=$ total variable cost at the point of maximum profits before contributions;

$k =$ constant overhead cost;

$$\frac{R-\left(\dfrac{1-\dfrac{1}{\epsilon}}{\sigma+1}\right)R}{(1+r)^{\left(\dfrac{1-\dfrac{1}{\epsilon}}{\sigma+\dfrac{1}{\epsilon}}\right)}} =$$ Total receipts less total variable costs at point of maximum profits after contributions;

$P =$ Net income (profits) before contributions.

The following observations can be made from equation (11):

(a) As σ increases, the quantity $(1+r)^{\left(\dfrac{1-\dfrac{1}{\epsilon}}{\sigma+\dfrac{1}{\epsilon}}\right)}$ decreases faster than the decrease in any other quantity containing σ.

(b) As ϵ decreases, the quantity $(1+r)^{\left(\dfrac{1-\dfrac{1}{\epsilon}}{\sigma+\dfrac{1}{\epsilon}}\right)}$ decreases faster than the decrease in any other quantity containing ϵ.

Hence, the greater σ, namely, the steeper the slope of the marginal cost curve, the smaller will be the relative decrease in profits as a result of contribution costs. Also, the smaller ϵ, that is, the less elastic the demand, the smaller will be the relative decrease in profits as a result of contribution costs.[5]

Given the above conditions and using equation (11), where R equals 100, k is such that maximum profits before the imposition of the contribution payments are 10 per cent of total receipts (value of product), and the contribution rate r is four per cent. Table I shows, among other things, the relative decrease in profits associated with various sloped marginal cost functions, while Table II indicates the same factors assuming various elasticities of demands.

From Tables I and II it is apparent that with a given relationship between profits and the value of product or receipts, and with a given demand function, the steeper the marginal cost curve, that is, the greater the rate of increase in variable costs with increases in output, the less will contribution payments act as an incentive to stabilize. In addition, with a given marginal cost curve, the more

5. In the case of monopoly, and when dealing with demand functions of constant elasticity, it must be assumed that the elasticity is always greater than unity, otherwise the solution is too far removed from reality even for theoretical purposes.
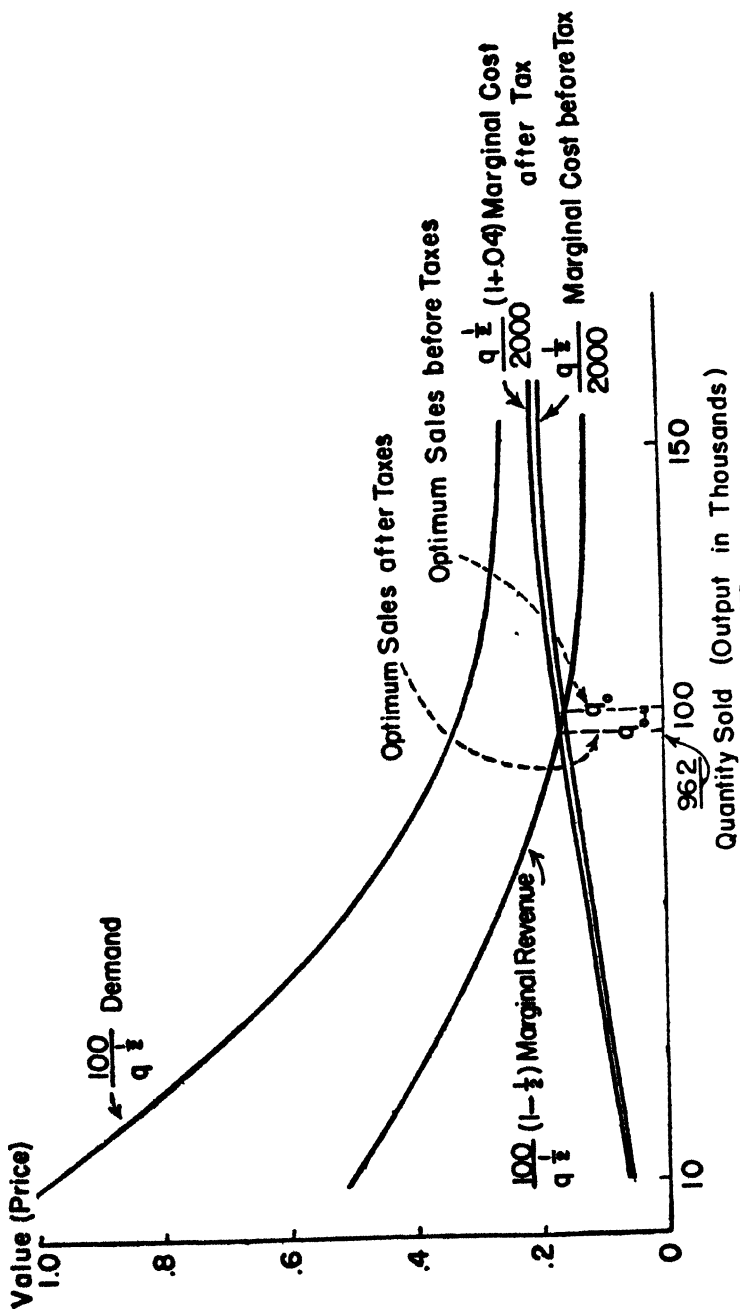
FIGURE III

DEMAND, MARGINAL REVENUE, MARGINAL COST, AND MAXIMUM AND MAXIMUM PROFIT OUTPUT BEFORE AND AFTER
CONTRIBUTION COSTS (TAX)

TABLE I

RELATIVE DECREASES IN PROFITS AND IMPORTANCE OF WAGES TO VALUE
OF PRODUCT (RECEIPTS) IN HYPOTHETICAL COMPANIES
UNDER VARIOUS CONDITIONS OF MARGINAL COST

Elasticity of Demand $\left[\dfrac{100}{q^{1/e}}\right]$ (e)	Marginal Cost $\left[\dfrac{q^{\sigma}}{2000}\right]$ (σ)	Profits as Per Cent of Value of Product Before Contributions	Per Cent Decrease in Profits if Price and Quantity Remain Same	Per Cent Decrease in Profits After Forward Shifting	Per Cent of Relative Decrease in Profits Shifted Forward	Wages as a Per Cent of Value of Product
2	⅛	10.00	17.78	17.16	3.47	44.44
2	¼	10.00	16.00	15 49	3.19	40.00
2	½	10.00	13.33	12.95	2.85	33.33
2	1	10.00	10.00	9.74	2.60	25.00
2	2	10.00	6.67	6.51	2.35	16.67
2	3	10.00	5.00	4.90	2.00	12.50

TABLE II

RELATIVE DECREASES IN PROFITS AND IMPORTANCE OF WAGES TO VALUE
OF PRODUCT (RECEIPTS) IN HYPOTHETICAL COMPANIES
UNDER VARIOUS CONDITIONS OF DEMAND

Marginal Cost $\left[\dfrac{q^{\sigma}}{2000}\right]$ (σ)	Elasticity of Demand $\left[\dfrac{100}{q^{1/e}}\right]$ (e)	Profits as Per Cent of Value of Product Before Contributions	Per Cent Decrease in Profits if Price and Quantity Remain Same	Per Cent Decrease in Profits After Forward Shifting	Per Cent of Relative Decrease in Profits Shifted Forward	Wages as Per Cent of Value of Product
¼	1.5	10.0	10.64	10.38	2.44	26.67
¼	2.0	10.0	16.00	15.49	3.19	40.00
¼	2.5	10.0	19.20	18.49	3.70	48.00
2	1.5	10.0	4.44	4.34	2.23	11.11
2	2.0	10.0	6.67	6.51	2.35	16.67
2	2.5	10.0	8.00	7.81	2.38	20.00

the elasticity of demand approaches one (elasticity always being greater than one), that is, the less total receipts increase with decreases in price, the less will contribution payments act as an incentive to stabilize.

Under the above conditions, the statement can be made that the greater the relative importance of wages to total cost at the point of maximum profits, the greater will profits be decreased relatively as a result of contribution costs,[6] and hence the greater

6. Cf. Harris, op. cit., p. 362.

will be the importance of r as an incentive towards stabilization (see the last column in Tables I and II). Figure III illustrates graphically the quantity which renders profits a maximum before and after contribution costs under one set of assumptions (see Table I).

SUMMARY

The burden of the contribution costs, in the case of monopolistic competition, cannot be shifted completely forward. The net income of the monopolist will always be decreased as a result of contribution costs. If the monopolist attempts to shift the contribution costs forward, his net income will be decreased less than if he does not attempt to do so. Further, if the monopolist attempts to pass on the increased cost through higher prices, his net income will be decreased by more than the amount of the contribution costs paid to the state; yet the decrease in his net income will be less than it would have been if he had not increased price.

Assuming a constant positively-sloped marginal cost function (marginal cost consisting only of wage cost), constant elasticity of demand and the existence of a given rate of profit before contribution payments are made, the steeper the slope of the marginal cost, the smaller will be the increased cost as a per cent of value of product (sales) and the smaller will be the relative decrease in the value of the product; hence the smaller will be the relative decrease in profits, and the smaller will be the relative burden (loss in net income) shifted forward.

In addition, the less elastic the demand (always being greater than one), the smaller will be the increased cost as a per cent of value of product and the smaller will be the relative decrease in the value of the product; hence the smaller will be the relative decrease in profits and the smaller will be the relative burden shifted forward. Conversely, the more elastic the demand, all other things being equal, and the flatter the marginal cost curve, all other things being equal, the more will profits be decreased relatively as a result of the unemployment compensation contribution, despite the fact that a greater per cent of the burden is shifted forward. Therefore, those monopolists with a relatively flat marginal cost function (usually associated with relatively little capital equipment) will have a greater incentive to stabilize than the monopolists with a relatively steep marginal cost curve (usually asso-

ciated with relatively large quantities of capital equipment). Furthermore, the monopolists with a relatively elastic demand will have a greater incentive to stabilize than will the monopolists with a relatively less elastic demand.

It should be mentioned, however, that unless the assumption is made that an equal rate of profit existed for all producers before the contribution payments were made, the relative decrease in profits may be greater or smaller from the sale of a commodity with a relatively elastic demand as compared with that from the sale of a commodity with a relatively less elastic demand. Likewise, unless the above assumptions are made, one cannot say anything about the relative decrease in profit from the sale of a commodity whose marginal cost function is relatively steep as compared with the sale of a commodity whose marginal cost function is relatively less steep[7].

Under many state unemployment compensation laws, a schedule of differential rates exists. Monopolists who can stabilize their employment will pay contributions at a lower rate and thus leave their net income decreased less with price increasing less and quantity sold decreasing less.

<div align="right">SAM ARNOLD.</div>

OHIO STATE UNIVERSITY

[7] The following table is an example of this point:

THE PERCENTAGE DECREASE IN PROFITS AFTER FORWARD SHIFTING IN HYPOTHETICAL COMPANIES UNDER VARIOUS CONDITIONS OF MARGINAL COST AND DEMAND, AND PROFITS BEFORE CONTRIBUTION PAYMENTS

Marginal Cost $\left[\frac{q^\sigma}{2000}\right]$ (σ)	Demand Function $\left[\frac{100}{q^{1/\epsilon}}\right]$ (ϵ)	Per Cent Profits Are to Value of Product Before Contribution Payments	Percentage Decrease in Profits After Forward Shifting
2	1.5	3.0	14.38
2	1.5	10.0	4.32
¼	1.5	10.0	10.38
¼	2.5	10.0	18.49
¼	2.5	25.0	7.40

From the above table, it follows that with a given rate of profit, the flatter the marginal cost curve, the greater the relative decrease in profits, and also, the more elastic the demand, the greater the relative decrease in profits. The opposite is true, however, if the profits are sufficiently greater in the case of the more elastic demand and if the profits are sufficiently smaller n the case of the relatively flatter marginal cost curve.

THE VOLUME OF FOREIGN TRADE AND THE LEVEL OF INCOME[1]

I. THE PROBLEMS

In analyzing the effects of foreign trade on the level of national income, the foreign trade multiplier approach has treated exports as income-creating and imports as leakages. As a result, only an export surplus is usually considered to be income-creating in the final analysis. Balanced changes in the total volume of foreign trade have, in this analysis, no effect on national income, since the income-creating effects of exports are exactly offset by the income-destroying effects of imports. Thus Professor Machlup writes: "If . . . imports (paid out of received incomes) jumped to the volume of exports without any lag at all, there would be no expansionary force in our . . . exports."[2] For foreign trade to have any expansionary effect there must be at least a temporary export surplus.

This article has three major purposes. First, a situation is constructed in which a balanced increase in trade is income-creating, and the conditions under which this is true are worked out. Second, it is shown that in so far as a balanced increase or decrease

1. I wish to thank Professors Schumpeter, Haberler and Samuelson, and Mr. Metzler and Mr. Downing for their criticisms of the earlier drafts of this article.

2. Fritz Machlup, International Trade and the National Income Multiplier, pp. 31, 32. In a footnote on page 31, Professor Machlup qualifies this statement in five ways, none of which is precisely what we are going to argue. On pp. 187ff. there are two sections on "Parallel Investment Expansions with Balanced Trade" and "Investments Increased in A and B in a Proportion which Equalizes Induced Imports." These cases are more similar to the discussion which follows. There are other instances in Haberler (Prosperity and Depression, Part III), Machlup (op. cit.), Robertson (Economic Journal, June, 1939), Harrod (International Economics, Rev. Ed.) which qualify the statement that only export surpluses are considered. Another case discussed by Mosak is quoted below. None of these discussions, however, refers to the precise points which we wish to make.

in foreign trade does change income, it is equivalent to a shift in the (domestic) spending function,[3] and the result reached therefore independent of the marginal propensities to import involved. Third, the reasoning is applied to certain likely postwar situations, specifically, the question of a permanent dollar shortage.

II. The Assumptions

Consider the following example. Suppose we have two countries, both in the midst of a depression. All resources are thus underemployed. Supplies of all commodities can expand in both countries without price changes, and national income in money and real terms will, under these conditions, move together. Interest rates are kept stable by Central Bank action, so that the supply of credit is perfectly elastic at the given rates of interest, and are assumed to be the same in both countries, in order to eliminate capital movements due to interest rate differentials. Such capital movements are extraneous to the problem which we wish to investigate. The exchange rate between the two countries is kept stable by rigid exchange controls, so that short-term capital movements or action by exchange equalization accounts are unnecessary. Except when otherwise stated, these will be the assumptions throughout. To simplify the discussion even more, it is assumed for the time being (and unnecessarily for the final outcome) that both countries have foreign trade monopolies, even though within each country the usual capitalistic institutions exist.

Suppose further, not only that the exchange rate is maintained by strict allocation of the available foreign exchange, but that both countries are willing to spend the proceeds from their exports on imports, that is, that neither country wishes, or can afford, to accumulate foreign exchange reserves.

Now both countries try to get out of the depression. They do not wish to incur budget deficits. Neither do they wish to increase the pressure on their foreign exchange control by expanding at home without making certain that the other country will expand also. If both countries negotiate with each other, they might come

3. The term "spending function" is used here instead of "consumption function," because we make not only consumption but also induced investment a function of income. In this we follow Tinbergen, who in another context has pointed out that "the essential issue is the choice between hoarding and spending, rather than between saving and consuming." "Critical Remarks on some Business Cycle Theories," Econometrica, Vol. 10 (1942), p. 145.

to the following agreement. During successive periods each country will make available to its importer from the other country increasing amounts of foreign exchange in such a way that during each period both countries' imports and exports are kept equal. To make certain that any results we reach do not depend on any lags whatsoever, it is assumed that trade is balanced within each income propagation period.[4]

In other words, both countries agree to keep stable prices, to maintain the exchange rate, but to increase the volume of foreign trade without creating a balance in either country's favor. To maintain the foreign trade at this higher level, there is thus no provision made for borrowing and lending, which both countries deliberately avoid. Both countries stick to the bargain, either because they are equally strong or because they voluntarily forego possibilities of economic warfare.

Now, on the assumptions usually made by the foreign trade multiplier approach, it seems that such a simultaneous increase in imports and exports can have no effect on national income, since, in the language of that approach, the multiplicand is zero.[5] However, the argument developed below indicates that even without a net foreign balance there may be an income-generating effect of the increased total volume of foreign trade, if the basic assumptions are slightly altered.

4. See Fritz Machlup, "Period Analysis and Multiplier Theory," reprinted in Readings in Business Cycle Theory.
5. This is only a slight overstatement. See below, p. 309 for the concurring opinion of Haberler. Cf. also the following quotation from Norman S. Buchanan, International Investment and Domestic Welfare. pp. 214–215: "A pay-as-you-go basis, on the other hand, appears to be continually neutral in its effect upon national income, or nearly so. The absorption of imports would be a depressant upon income and employment in the United States according to the usually accepted analysis. But if the dollar exchange obtained from imports were almost immediately spent for goods to be exported abroad, there would be a corresponding (and almost immediate) stimulus to incomes and employment at home.

Hence as between a loan policy and a pay-as-you-go policy in their effects upon national income the main distinctions would appear to be these. A loan policy gives a protracted stimulus to employment and income as long as there is a net outflow of investment abroad. As Mr. D. H. Robertson once said, referring to Great Britain, 'For getting out of a slump there is nothing like a whiff of foreign investment.' But he might have added that the return flow of capital as interest and principal was a soporific. A pay-as-you-go policy, on the other hand, depresses incomes in one sector of the economy at the same time that it stimulates another. The tendencies probably about neutralise one another, except in special circumstances."

These changes in the assumptions seem both reasonable and important. The foreign trade multiplier analysis proper asserts that, given the spending function (i.e. the total average propensity to spend) an export surplus will have an expansionary effect, which then depends on the marginal propensity to import. My conclusion rests on the argument that *a balanced increase in the total volume of foreign trade will itself*, under the assumptions, *raise the total average propensity to spend.*[6]

While the assumptions are stated in political terms to indicate what might be involved in reality, I should like to point out that the conclusion that a balanced increase in the total volume of foreign trade will result in an upward shift of the spending function will be correct, in general, both in cases where the spending function had been artificially depressed before the balanced increase in trade, through exchange controls, quotas or other interferences, and those where the spending function can be considered to shift for any reason whatsoever.

The former is self-evident. The likelihood of autonomous shifts, though self-evident to me, has been questioned by some of my friends, and so the following reasons may be given in addition to those which will be given below in section V. It seems plausible that the consumption function and its extension, the spending function, will shift during a cycle and with the movements of income. One reason for this shift may be the changing expectations as to the future course of the cycle. Another is that changes in the income level are accompanied by changes in the income distribution, which "are very important for the propensity to consume of society as a whole, even from the short-run point of view."[7] I shall further argue in section V that changes in trade policy must

6. The word "total" is used to signify that we are speaking of the average propensity to spend on both imports and domestic goods. The "spending" function may be interpreted either so as to include spending on both imports and domestic goods or, alternatively, since exports are equal to imports, it can be thought of as including spending on domestic goods only, that is, goods for domestic use and exports of home-produced goods. This latter is the interpretation which will be used here.

7. G. Haberler, Prosperity and Depression, 1941 ed., p. 229. This is equally true of the marginal and the average propensity to consume.

If we relax the assumption of fixed prices we ought to add that price changes also must be presumed to lead to shifts in the spending function. The insistence of the classical economists that an equilibrium with full employment will be reached with flexible prices (including interest and wage rates) can be translated into Keynesian terms in this way.

themselves be expected to lead to shifts of the spending function which cannot be neglected and which are usually assumed away. I believe, therefore, that both the assumptions and the results reached are of far wider significance than is indicated by their "political" formulation, which might appear to restrict the reasoning to State trading. A theoretical framework which allows for shifts in the spending function is more general and more significant.

III. SYMBOLIC VERSION

The following algebraic and numerical version should clarify the precise meaning of our assertion, and should bring out clearly the difference between my assumptions and those usually made. Following Metzler,[8] we let income of each country (Y and Y', respectively) consist of domestic consumption (C and C'), domestic investment (I and I') and exports (E and E'). I assume the same time relationships as Metzler does. Consumption in each country depends on the income of the previous period and the marginal propensity to consume of that country. Similarly (induced) investment depends on the domestic income of the previous period and the marginal propensity to invest. Exports depend on the marginal propensity to import and the income of the previous period of the *other* country. In symbols, we have therefore, according to Metzler:

$$(1) \quad \begin{aligned} Y_t &= C_t + I_t + E_t \\ Y_t' &= C_t' + I_t' + E_t' \end{aligned} \quad \begin{aligned} & \text{for country I} \\ & \text{for country II} \end{aligned}$$

$$(2) \quad C_t = aY_{(t-1)} \quad ; \quad C_t' = a'Y_{(t-1)}'$$

$$(3) \quad I_t = \beta Y_{(t-1)} \quad ; \quad I_t' = \beta'Y_{(t-1)}'$$

$$(4) \quad E_t = \gamma'(Y_{(t-1)}') \quad ; \quad E_t' = \gamma Y_{(t-1)}$$

where a, β, and γ stand for the marginal propensities to consume, to invest, and to import, respectively, and where the absence or presence of primes indicates that we are dealing with country I or II, respectively.

Now our change of assumptions leaves equations 1, 2 and 3 unchanged. But equation 4 is altered. For, as each country will purchase the same increased amount from the other, equation 4 will be changed to

$$(4a) \quad E_t = \gamma'Y_{(t-1)}' + r_t' = \text{constant} = E_t' = \gamma Y_{(t-1)} + r.$$

8. Lloyd A. Metzler, "Underemployment in International Trade," Econometrica, Vol. 10 (1942); "The Transfer Problem Reconsidered," Journal of Political Economy, Vol. 50 (1942).

The first parts of these expressions are induced imports, which depend on income and the marginal propensity to import. But the second parts, r and r' respectively, are the imports which are needed to maintain the volume of trade at the agreed higher level. It would be easy to change this equation appropriately if the original assumption were, for example, that foreign trade is raised periodically to a new and higher level in a balanced fashion.

This symbolic expression makes it immediately clear that a balanced increase in trade will raise income precisely as any autonomous increase in either consumption or investment at home. Equation (1) is changed to

(1a) $Y_t = C_t + I_t +$ constant; $Y_t' - C_t' + I_t' +$ constant. The marginal propensities to import thus drop out, and the final effect of the balanced increase in trade will depend exclusively on the domestic marginal propensities to spend.

Rather than complicate the argument by changing the original assumptions, let us consider a numerical example adapted from the tables made familiar by Machlup and Metzler. I choose Metzler's set-up for reasons which will be given presently.

Metzler has shown that the effect of international trade on income will depend on, first, the stability conditions, and, second, on the monetary policies pursued by the individual countries. Both our countries are assumed to be stable in isolation, in order to show that the income-creating effects of a balanced increase in foreign trade do not depend on any instabilities. And both countries allow their income to be affected immediately. The agreed increase in trade, which is then maintained at this higher level, is assumed to be 10 units. Table I is read precisely as is Metzler's table from which it is adapted.[9] It differs from his in two respects, both differences being dictated by the different assumptions. It contains two additional columns (2 and 8). The income of both countries rises immediately, as both countries let their incomes be affected directly. The income of country I is found by adding columns 1, 4, 7 and 8; the income of country II consists of columns 6, 9, 2 and 3. But since columns 7+8 equal columns 2+3, and both by assumption equal 10, we could add immediately columns 9, 2, 3 and 4 to find the income of country I, and columns 6, 7, 8 and 1 to find the income of country II. This is simply the numerical expression for the assertion that a balanced increase in foreign

9. Metzler, Transfer Problem, op. cit., Table 1, p. 401.

trade is equivalent to an upward shift of the domestic spending function.[1]

It may be worth while to add a word here about one difference in the construction of the tables presented by Metzler and Machlup. Both authors start with the same identity: the marginal propensities to consume, to import, to invest and to save add up to unity. Machlup — legitimately for his purpose — neglects the marginal propensity to invest. Machlup states explicitly the marginal propensities to save and to import, while Metzler states explicitly the marginal propensities to consume and to import and to invest. Formally the two approaches are identical, for the missing marginal propensity can always be found by subtracting the stated propensities from 1.

Even though both ways of stating the problem simply say that an increase in income will either be spent or not, and that that part which is spent may be spent either at home or abroad, they hide a difference in assumption which is important from our standpoint. In Machlup's case, an increase in the marginal propensity to import is always a leakage, because it implies a decrease in the (not specifically stated) marginal propensity to consume domestic goods. With Metzler, an increase in the marginal propensity to import implies a decrease in the (not specifically stated) marginal propensity to save. Metzler's procedure is better adapted to our purpose while Machlup's procedure seems ideally adapted for a study of the foreign trade multiplier proper.[2] The assumption that imports are a leakage is no more than an assumption. It seems no more legitimate or self-evident than the assumption which Metzler and, following Metzler, this article make.

The equilibrium condition is evidently that the average pro-

1. It would be easy to construct a table which shows the effects of both a compensated increase in foreign trade and the leakages of the foreign multiplier proper. The "realistic" assumptions would be, for example, that our countries agree to increase foreign trade by a minimum amount, say 10 units, in a balanced fashion, but to permit net balances over and above this amount. Numerically this would be done by keeping columns 3 and 8 constant at 10 units, but letting columns 2 and 7 change depending on the marginal propensities to import. Income of country I would then consist of columns 1, 4, 7, 8 = columns 1, 4, 7, 3. And income of country II would consist of columns 6, 9, 2, 3 = 2, 6, 8, 9. In this case the marginal propensities to import would have to be considered for the final outcome.

2. This difference in stating the assumptions may lead a careless reader of Machlup's work to suppose that he is dealing with unstable cases, when he is, of course, not doing anything of the kind.

TABLE I

INCOME EFFECTS OF BALANCED INCREASE IN TRADE OF 10 UNITS. BOTH COUNTRIES ARE STABLE IN ISOLATION. BOTH COUNTRIES LET THEIR INCOMES BE AFFECTED DIRECTLY.

	Country I	Country II
Marginal Propensity to Consume Domestic Goods	0.2	0.3
Marginal Propensity to Consume Foreign Goods	0.2	0.1
Marginal Propensity to Invest	0.1	0.1

Period t	Country I					Country II				
	Consumption of Domestic Goods C	Induced Consumption of Foreign Goods	Auton. Consumption of Foreign Goods	Investment I	Income Y	Consumption of Domestic Goods C'	Induced Consumption of Foreign Goods	Auton. Consumption of Foreign Goods	Investment I'	Income Y'
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
1	...	...	10	...	10	...	...	10	...	10
2	2	2	8	1	13	3	1	9	1	14
3	2.6	2.6	7.4	1.3	13.9	4.2	1.4	8.6	1.4	15.66
4	2.78	10	...	1.39	14.17	4.68	10	...	1.56	16.24
5	2.834	10	...	1.417	14.251	4.872	10	...	1.624	16.496
:	:	:	:	:	:	:	:	:	:	:
	2.8572	10		1.429	14.286	5.0006	10		1.6667	16.667

pensity to spend is equal to one. The total volume of exports (which equals the total volume of imports) may be imputed to domestic consumption and/or investment. In neither case are imports a "leakage." How it is to be treated in any particular historical situation is simply a matter of fact. To repeat: what matters is only that there is a net increase in the average propensity to spend. It is clear, for example, that an import surplus will lead to a fall in domestic incomes only if it has led to a reduction of spending on domestic goods. But it cannot lead to a fall in income and employment if it has led to a reduction of savings instead, or if it is financed by additional foreign borrowing. There must, then, be a net stimulating effect if the imports are paid for by exports which themselves are financed with additional money, so that there is no offsetting change in home consumption and home investment.[3]

IV. GRAPHIC PRESENTATION

The familiar graph of the consumption function can be used to picture the reasoning of the preceding section. Starting with income Y_1 (Figure I), the income Y_2 could evidently be reached and maintained either with a stream of investments AB, if the consumption function is CC', or without any investments at all, if the consumption function is KK', or if it has shifted to this higher position. An export surplus can be treated as part of investments. For our discussion it is, therefore, important to remember that an autonomous increase in consumption — i.e. an increase in consumption which is not due to, and is independent of, a change in income — will have the same effect on income as an autonomous increase in investments.[4]

Applying the reasoning to our particular problem we can say that the increased volume of foreign trade is equivalent to an increase in the average propensity to import (which previously had been held in check by exchange control) with no corresponding offsetting decline in the average propensity to consume domestic goods. It has thus raised the total average propensity to consume. And, as we have seen, since exports are assumed to be equal to

3. Below we will discuss in greater detail the reverse case, namely, that an export surplus cannot automatically be considered to have an expansionary effect. See below, p. 296.

4. While autonomous changes in consumption are represented by shifts in the function, induced changes are represented by movements along the function.

imports, we can consider consumption to consist only of domestic goods. If before the trade agreement income was Y_1 and the consumption function CC', the balanced increase in trade will shift the consumption function to KK'. This will raise income just as much as an export surplus of $MN = AB$ would have done. *Furthermore, the increase in the average propensity to consume will have secondary effects just as an export surplus would have had.* The graph immediately shows that Y_2 will be greater than income Y_1 by more than the volume of foreign trade, and also that this will depend only on the domestic marginal propensities.[5] The marginal propensity to consume has been assumed to be equal for CC' and KK'.[6] The case in which the total volume of compensated imports is assumed to have added to domestic investment rather than domestic consumption is shown on Figure I as follows. Starting with income Y_1, MN is now interpreted as being new net investments which are financed with additional credits.[7] The final adjustment would, of course, be income Y_2 as before, even though the consumption function has not shifted upward. However, investment $MN = AB$ could consist entirely of imported goods paid for by additional exports. Consequently, with the consumption function unchanged, the total volume of exports must in this case too be credited with the stimulating effect on income and employment.

V. The Case of a Tariff

My thesis is thus that under the most probable assumptions, a simultaneous rise (fall) of imports and exports will lead to an increase (decrease) in the average propensity to spend on consumption and/or on investments. Any measures to expand

5. It has been shown in the preceding section that the marginal propensity to import is irrelevant for the final outcome under our assumptions. On the graph we also omit the repercussions on domestic investment which were shown in Table I. This involves no loss of generality.

6. It might be mentioned in passing that this well-known graph, which as far as I know is due to Professor Samuelson, can also be used to illustrate the underconsumption theory as formulated by P. M. Sweezy (Theory of Capitalist Development, p. 183). If I understand Mr. Sweezy's theory correctly, it can be most simply described by stating that the full employment income has, because of capital accumulation, a tendency to shift to the right faster than consumption shifts upwards.

7. Evidently, if the net new investment had been financed by additional savings rather than additional credit, there would have been no *net* increase in spending, but only a redirection of the amounts already spent before.

(contract) the total volume of international trade will lead to an increased (decreased) income and employment, provided the average propensity to consume and/or invest is increased (decreased). A lowered tariff, for example, will lead to increased income and employment everywhere, provided the added opportunities for consumption are realized.[8]

Under what conditions does a simultaneous increase (decrease) in imports and exports actually involve a rise (fall) in the total average propensity to consume? Since we have so far considered the case of a simultaneous *increase* in foreign trade, we shall now discuss the reverse problem. Specifically: what will be the short-run effects on income and employment of a new tariff?

It is interesting to note here a certain similarity between the classical and multiplier approaches. Because the classical approach assumes full, or at least fixed, employment (in the long run), a new tariff will simply induce people to spend money at home instead of abroad, substituting their expenditure for the spending of the foreigners. There is (in the long run) a complete shift from foreign to home-produced goods, and the total average propensity to consume remains unchanged. So, of course, does employment by assumption.

The usual multiplier analysis ordinarily simply implies that a newly imposed tariff will leave the total average propensity to consume unchanged. For an export surplus due to a new tariff will have an expansionary effect only if there is at least a partial shift from the consumption of foreign to the consumption of home-produced goods. True, the foreign trade multiplier effect depends on the fact that the tariff is expected to produce at least a temporary export surplus.[9] But unless there is some shift of consumption

8. In this section we shall from now on make the simplifying assumption that all imports are for consumption. Since autonomous changes in consumption and investment have identical effects on income and employment, this convenient simplification involves no loss of generality. The reader is reminded that we are throughout the article making the assumptions outlined above in section II, except when otherwise stated. These include, specifically, unemployment, elastic credit supply, constant prices.

9. That Machlupis quite aware of the case which we are discussing is shown by the following quotatior, which also indicates why he did not analyze it: "We mentioned at the outset of our demonstration of the *modus operandi* of the foreign-trade multiplier that the case of an autonomous import reduction, under the chosen set of assumptions, would be fully analogous to that of the autonomous export increase, which is the standard case of the analysis. That an export increase and an import reduction, thrown together under the head-

from previously imported to domestic goods, there can be no net stimulating effect on employment and income; at most this export surplus would have no effect at all. Foreigners and natives would consume the goods of the first nation as before, but natives would buy less foreign goods. Employment depends on total demand whether by natives or by foreigners. This total demand is unchanged, unless there is a shift towards the consumption of home-produced goods. Total employment would thus remain unchanged.

Eventually there might be a net stimulating effect even in this case in which there is no shift. But such a stimulating effect would in such a case come about in the most orthodox fashion from an export surplus *via* a gold inflow and/or an easing of the balance of payments situation to an easier monetary situation at home, and finally *via* a lowered interest rate to increased investments. This indirect effect is not what we are interested in, however, and we have specifically excluded it by our assumption of constant interest rates. Nor is this the case discussed by the foreign trade multiplier analysis proper.

We can go further. If exports as well as imports fall immediately as a result of the tariff, there will be an immediate fall in total demand, and therefore employment, unless there is a shift of sufficient magnitude towards the consumption of home-produced goods. This is true even if the tariff leads to an export surplus, as exports fall more slowly than imports. Only if there is a shift of consumption away from imported goods towards home-produced goods can there be a stimulating effect of an export surplus. We can therefore state in general:

(a) If the export surplus comes about through a fall in imports with exports unchanged, it will have a stimulating effect only if the average propensity to consume domestic goods is increased. Otherwise the increase in the average propensity to invest (as exemplified by the export surplus) will be exactly offset by the decline in the total average propensity to consume.

ing of 'improvements of the trade balance,' are equivalent in most of their effects is often an absolutely unwarranted contention; but under the strict regime of the chosen assumptions the equivalence holds. . . . After all . . . the assumptions include unchanged commodity prices, unchanged interest rates, unchanged foreign-exchange rates and, of course, absence of further 'autonomous' changes in trade such as would be implied in retaliatory import restrictions by the adversely affected country. This is just too much to neglect. But for the purposes of a clean analysis we stick to our assumptions." Op. cit., pp. 58–60.

(b) If the export surplus is due to the fact that imports are falling faster than exports, it will have an expansionary effect only if the average propensity to consume domestic goods rises faster than exports fall. Otherwise the total average propensity to consume plus the average propensity to invest will be smaller than they were before the export surplus came about.

(c) Only if the export surplus arises because exports increase with imports increasing less or remaining constant is it unnecessary for the average propensity to consume domestic goods to increase in order to achieve an expansionary effect.

In the classical case of full or fixed employment, a tariff reduction evidently cannot lead to an increase in the total average propensity to consume. We have, however, assumed unemployment. Consequently a raised (lowered) tariff need not simply lead to a substitution of domestic for foreign goods (or vice versa). It may, and probably will, lead to additional savings in the case of a raised tariff, because consumers, faced with a restriction of the range of goods available to them are likely to restrict their total consumption. Similarly, the case of a lowered tariff may lead to additional foreign demand financed either out of idle funds or by bank credits in the manner indicated below (section VI).

Now, if a tariff is imposed there will be one or more of three effects. First, there will be a tendency for exports to fall later than imports. This is the foreign trade multiplier proper. Second, as prices of imported goods tend to rise, natives will substitute domestic goods for imports. That is, the average propensity to consume domestic goods will rise simultaneously with a fall of the average propensity to consume foreign goods. Third, there will be a substitution of savings for imports. That is, natives might decide to spend less altogether.

If we now disregard for the time being the foreign trade multiplier effect proper by sticking to our original assumption that foreign trade moves in a balanced fashion, the last two substitutions give a clue to the various possibilities.

(a) A balanced decrease of imports and exports, due, for example, to a new tariff, will lead to an increase in employment and income if natives substitute domestic goods for foreign goods fully, and if, in addition, they substitute consumption of domestic goods for savings.

(b) A simultaneous decrease of imports and exports will leave

total employment and income unchanged if natives substitute domestic goods for foreign goods, and if there is no substitution between savings and foreign and domestic goods.

(c) A simultaneous decrease of imports and exports will lead to a fall of total employment and income if natives fail to substitute domestic goods for foreign goods completely, and if they instead substitute savings in part for imports.

To put it somewhat differently, the effect of a fall of total foreign trade will depend on whether natives increase their average propensity to consume sufficiently to offset the fall in the foreigners' propensity to consume the first country's goods, or whether this increase is more or less sufficient for this effect.

Is there any reason for believing one of these three possibilities more probable than the others? On the one hand, we can argue that there must be a tendency for the third possibility to arise. Therefore the total average propensity to consume will tend to fall when foreign trade decreases in a balanced fashion. But there may be one offsetting tendency. Both Keynesians and non-Keynesians generally agree that real savings will diminish as real income falls. We have therefore contradictory tendencies, and unless we can say more about them than we have done so far, the final outcome must remain in doubt.

Now there can be no doubt that a tariff will lead immediately to a reduction of real income,[1] if it leads to higher import prices and to a substitution of domestic for imported goods under unfavorable circumstances. On the other hand, it might be argued that the case for a reduction of real income as the result of a tariff is clear-cut only if there is full or fixed employment. If there is under-employment it is certainly possible that a substitution of domestic for foreign goods may leave real income unchanged. There are two presumptions against this reasoning. The first is that export industries are more efficient than domestic industries, and that therefore a shift in employment will lead to lowered efficiency everywhere. The second is that consumers preferred

---

1. Except for this paragraph we could use the terms real and money income interchangeably, because prices were assumed to be constant. If prices change, real and money income may move in opposite directions, and we have therefore to specify whether we talk about money or real income. It is the purpose of the present paragraphs to determine whether a tariff is likely to increase or decrease money income without moving real income in the same direction — or vice versa.

imported goods or else they would have bought domestic goods in the first place. The first presumption loses much force when unemployment exists. The assumption that the production of all goods can be expanded without price changes seems legitimate. As to the second presumption, we may assume — less legitimately — that natives are indifferent as to whether they consume domestic or foreign goods.[2]

The point I wish to make is this: real savings will fall if real income falls. But with underemployment of all resources real income will fall as a result of a tariff only if there is less than a complete substitution of domestic goods for imports, that is, only if the total average propensity to consume falls. Consequently, the decrease in real savings will slow up the fall in real income and employment if there was less than a complete substitution of domestic for foreign goods. But the fall in savings cannot entirely counteract the absence of complete substitution.

Under the assumption originally adopted it seems that the most probable case will be the one in which total income will fall with a simultaneous decrease of imports and exports. Retaliatory tariffs will make this outcome a certainty, and so will exchange controls. But even under less stringent assumptions than we have made, it seems probable that there will be some substitution of savings for imports, and that therefore the most probable result of a tariff will be to reduce the total average propensity to consume. It follows that *mutatis mutandis* a simultaneous increase in imports and exports will, *via* an increase in the total average propensity to consume, raise income and employment everywhere.[3]

The discussion of the various possibilities of substitution indicates one way in which the foregoing discussion may be joined

2. Under competitive conditions and in equilibrium this is, of course, not only legitimate but necessary.

3. For the purposes of this article it is legitimate to speak as if an increase in employment and in income were the same thing. If we speak of employment of all factors, and if we excluded foreign borrowing, then more income can be produced only if more factors of production are used. But if we speak of employment of any one factor, for example labor, then we should have to take into account also the factor proportion necessary for the production of foreign and domestic goods.

I should like to point out that the reasoning here applies in the short run. If it is correct, the foreign trade multiplier cannot be used to defend protection as income-creating even in the short run without heavy qualifications. No reputable economist has, of course, ever used the foreign trade multiplier to defend tariffs.
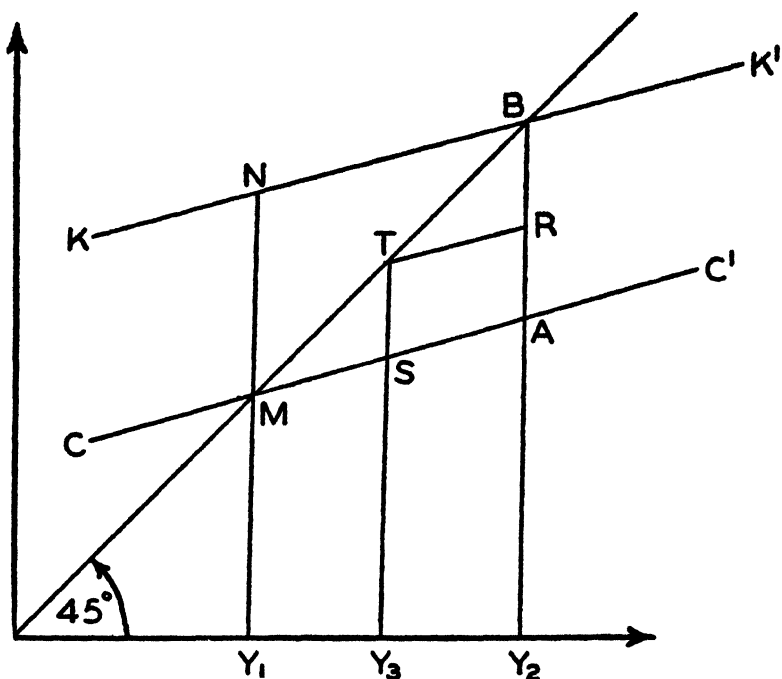
FIGURE I

to the usual multiplier analysis. We drop for the moment the assumption that imports and exports decrease together and assume that exports decrease less than imports as the result of the tariff. Then we get two contradictory effects. On the one hand, the total average propensity to consume declines. On the other hand, the export surplus is equivalent to an increase in the average propensity to invest and will thus have an expansionary effect. Whether the net effect will be expansionary or not depends on whether the average propensity to consume falls by more or less than the average propensity to invest (as exemplified by the export surplus) rises. The case in which the fall of the average propensity to consume more than counterbalances the effects of the export surplus cannot be dismissed as unrealistic or unimportant.

This is pictured on Figure I as follows. Let the original consumption function be KK'. The tariff depresses the consumption function to CC', but leads simulatneously to an export surplus of $AR=ST$. The final income reached under these conditions will have fallen from Y_2 to Y_3.

VI. Analysis in Terms of Money Flows

The inherent reasonableness of the above argument can perhaps be made clearer by analyzing briefly the effects which the change in the total average propensity has on the amount and flow of money. Since we have assumed the supply of credit to be perfectly elastic in both countries, it seems logical to assume further that in both countries the banking system is willing to finance the increased foreign trade and thus to expand credit in the most time-honored textbook fashion. If we wish to be more realistic we could assume government guarantees of competitive traders, instead of a foreign trade monopoly.

In both countries the financing of the additional exports will lead to the following balance-sheet changes in the commercial banks. First, deposits and commercial paper of the banks go up, the additional deposits belonging to the additional exporters. The financing of the exporters was necessary to enable them to pay manufacturers. Since there was underemployment in both countries, and since the exporters did not receive money at the expense of anyone else, production, employment and (money and real) income have expanded in both countries *pari passu.*

The further steps are that while total assets and deposits of the banks remain unchanged at the new higher level, the ownership of the deposits changes as the newly-created deposits move from the exporters to the manufacturers, to the workers, to domestic consumers, and so on. Now, the additional money in both countries is available to pay for the additional imports. The international offsetting of the accounts means only a change of ownership of the deposits within each country. It does not mean the disappearance of the newly created money.

It can be seen that all steps from the financing of the additional exports on can take place simultaneously in both countries — all, that is, except the final repayment of the loans. Because of the trade agreement which we have postulated above, the exports can take place in both countries simultaneously without a preceding stimulus from the other country. The expansionary effect of the raised level of foreign trade depends, in terms of the money-flow analysis, solely on the fact that it was financed with additional money rather than out of additional saving. It does not depend on timing. The creation of the additional deposits is the conse-

quence (or the mirror image) of the increased average propensity to consume.

We could assume that the trade agreement provides for an increasing stream of trade between the two countries in successive periods without a change in the trade balance. But this is really assuming more for our problem than is necessary. The newly created deposits will disappear from the economic system only if trade recedes to the *previous* level. As long as there is a constant higher (or even increasing) level of trade, traders need a constant (or even increasing) amount of newly created deposits. Thus the credits created to finance the increased volume of trade turn out to be a fund which is as permanent as the increased volume of trade it has to finance. This fund is, to be sure, a revolving fund, but it must exist to the extent of the *volume* of exports, not simply to the extent of any export surplus. As long as trade is really maintained at the higher level, there is no reason why the amount of money should ever decrease.

We may conclude, therefore, that in this case, and barring secondary repercussions, the income of both countries will be raised by the amount of foreign trade as long as this particular level of foreign trade continues to exist. It can also be seen from the discussion that as long as the total volume of trade between the two countries increases (and as long as there is underemployment in both countries), the income of the two countries will keep on increasing, both in real and money terms. If the volume of trade reaches a constant level, the income of both countries will stop increasing, but it certainly will not fall.

VII. Comparison with Self-Help Schemes for the Unemployed

The foregoing analysis bears some very obvious resemblance to the many self-help schemes for the unemployed which were so common during the great depression. One of these is discussed by Mrs. Robinson. She points out that "provided he does not sacrifice his right to the dole, an unemployed man does himself good, and on balance does others no harm, by occupying himself as usefully as he can."[4] Her argument is even more valid if the unemployed who are working for each other are financed with additional money rather than left to barter with each other. Her qualification that

4. Joan Robinson, *Essays* in the Theory of Employment, "Disguised Unemployment," pp. 82ff. The quotation is on page 99.

the dole must not be sacrificed is exactly parallel to the proposition that the average propensity to consume domestic goods must not decline when the average propensity to consume imported goods increases. The main difference between Mrs. Robinson's case and mine is the evidently rather secondary one that the unemployed working for each other live in different countries and have to be financed by different banking systems.

VIII. APPLICATION TO TWO CASES

The case analyzed here does not seem to be just a theoretical curiosity without practical implications. In order that the result indicated might be reached it was convenient to assume the existence of foreign trade monopolies. It is sufficient, however, that clearing agreements exist, with or without provisions for multilateral offsets, but with provisions for settling accounts at regular short intervals, say four times a year. Even less international "planning" than this is necessary to achieve the results outlined. We shall discuss two situations to which the present argument seems to apply.

Exchange Depreciation. During the depression of the 1930's, foreign trade of most European nations with each other declined. It does not matter in this context what precisely caused this decline. It remains a fact, and it is also a fact that the beggar-my-neighbor policies of the various countries were largely self-defeating. On the other hand, Mr. Harris, in his well-known study, found that trade between Great Britain and other paper-money countries increased after Great Britain had depreciated and other countries followed suit or joined the sterling block. He formulated one of his conclusions as follows: " . . . a country gains an export advantage if its currency depreciates; but though a study of export trade of paper countries in three groups of countries (classified according to currency conditions) reveals relative gains, the rise in the share of exports of paper countries taken by other paper countries suggests the importance of a rise of income and improved trade relations among paper countries,) and declines and new trade barriers in gold countries)."[5]

This result can easily be re-interpreted in the following manner. Because of expected balance-of-payments trouble, the average propensity to import is depressed by deliberate policy. Going off

5. Seymour Harris, *Exchange Depreciation*, p. 13.

the gold standard not only makes it possible for the countries involved to expand at home without worrying unduly about changes in impoʌts induced by rising incomes, that is, about the foreign trade multiplier effect proper; it also makes it possible to allow the average propensity to import to find its own level. A simultaneous increase in the average propensities to import (and thus of the average propensities to consume) of all countries concerned is in this situation not only possible but probable. It is the most reasonable way to explain the fact that trade between paper countries increased with simultaneous rises in incomes, but without either affecting exchange rates, or making exchange controls necessary, or without any capital movements.[6]

This point is perhaps worth discussing in somewhat greater detail. Mr. Metzler, in his path-breaking analysis, makes the following point: " . . . an increase in the average propensity of Y to consume Y' goods must increase total consumption and income in both countries. In the case of Y' this is obvious, since the increased demand of Y' goods has resulted in greater foreign investments. In the case of Y it means simply that the stimulating effects of a higher aggregate propensity to consume plus the secondary repercussions of higher income in Y' have exerted a more important influence on income of the former than the depressing consequence of increased imports."[7]   This conclusion must hold true *a fortiori*, if there is a simultaneous shift upwards in the average propensities to import in both countries, other things remaining equal. The leakages due to induced imports will, of course, still be present, and on our assumption of a parallel shift of the consumption function they will be equal to what they were before the shift. But the net income-creating effect of the increase in the volume of trade remains, as we have seen. The common sense of the proposition should be obvious. No one can deny the reality of an increased income with unchanged domestic investment, unchanged consumption of domestic goods, and an increased consumption of imported goods paid for by an increased production of goods for export financed by additional credit without foreign borrowing.

6. Some capital movements did, of course, take place. A simultaneous increase in investments in the countries concerned would equally account for the facts. The two explanations may thus be alternatives. But they can be, and probably usually are, complementary to each other.

7. Lloyd A. Metzler, "Underemployment Equilibrium in International Trade," *Econometrica*, Vol. 10 (1942), p. 108.

Perhaps the growth of bilateralism and of barter agreements in general, and theirattractiveness under depressed conditions, can also be partly explained and rationalized by analogous reasoning.

A permanent dollar shortage? If the reasoning developed above gives an additional explanation of the income effect of exchange depreciation, and thus adds perhaps to the understanding of the 1930's, it can also be applied to the problem which may confront the United States in the future. Mr. Kindleberger describes the likely postwar situation in the following terms: "The United States could import more finished goods at any level of production, can import more raw materials at higher levels of production, and might import more agricultural products to the extent it succeeds in moving factors of production already engaged in agriculture into industry. But these increased imports will raise money incomes abroad and will produce increased demands for American goods in excess of the original increases in American imports. The chronic shortage of dollars would remain, albeit at higher levels of real income throughout the world."[8]

Offhand, it seems as if Mr. Kindleberger explained a permanent shortage of dollars by assuming that foreign countries are unstable in Metzler's sense. However, this is not necessary for the outcome, and instability would be a condition imposed on top of the all-round increase in the average propensity to import American goods.

It is easy to imagine the existence of the following situation in the not too distant future. The marginal propensities of foreign countries to consume their own goods and to import are given, and do not change autonomously but only, if at all, in response to income changes. Domestic investments depend on the marginal efficiency of capital which in this context can be assumed to remain constant. Now, imports depend, on the one hand, on the domestic income level. But they also depend on the availability of foreign exchange, that is, on exports. The first of these relations is given

8. Charles P. Kindleberger, "Monetary Stabilization," in *Postwar Economic Problems*, edited by S. E. Harris, p. 387. Under our assumptions, money and real incomes would evidently rise together.

It should be added that the primary purpose of this article is theoretical, rather than to discuss precisely how serious the dollar shortage is apt to be. While most observers seem to believe such a shortage to be imminent, there has recently been one notable dissent by Lord Keynes ("The Balance of Payments of the United States," *Economic Journal*, June, 1946 pp. 172ff.) which appeared after the present manuscript was finished.

by the marginal propensity to import. The second relation can be better described by means of the average propensity to import.

We can now speak of induced-autonomous imports.[9] The imports are autonomous in the sense that they are independent of the level of income in the importing country. They are induced in that they depend on the level of exports, and thus on changes which occur in the other country. Therefore, even though any rise in income still leads to a smaller rise in *induced* imports — so that the country is stable in Metzler's sense — a rise in imports leads immediately to a shift in the average propensity to consume upwards to the extent of the increased exports.

Specifically, the situation may possibly exist in which foreigners prefer American goods to their own but will buy their own *faute de mieux*, and for the rest go without.[1] Their purchases of American goods are limited only by the amount of foreign exchange available to them, that is, by their exports plus such amounts as they can borrow. We may omit borrowing as extraneous to our problem. It follows that any increase in the American average propensity to consume will lead immediately to an increase in American and foreign incomes, even though American imports lead immediately to an equivalent increase in American exports. The effect is independent of surpluses of any kind, nor does it depend on instabilities as discussed by Samuelson and Metzler.

Table II summarizes this case. Country I corresponds to the United States, country II to the rest of the world. We need not assume any marginal propensity to import for the rest of the world, since we assume that they will buy automatically to the extent of the American purchases from them. Column 6 is therefore always equal to column 2. We assume again that both countries let their income be affected directly. And we start again with an autonomous increase in United States imports of 10 units. Income of country I (United States) is found by adding columns 1, 2, and 3. Similarly the income of the rest of the world is found by adding columns 5, 6 ($=2$) and 7.

The table indicates, that, if both the world and the United States are assumed to be stable in the Samuelson-Metzler sense,

9. This term is no more of a self-contradictory monstrosity than monopolistic competition.
1. The argument is strengthened if it is further assumed that foreign investments (and thus employment) depend on American exports.

TABLE II[a]

EFFECT OF A BALANCED INCREASE IN FOREIGN TRADE, IF COUNTRY II's IMPORTS DEPEND ON ITS EXPORTS, AND COUNTRY I's IMPORTS ON ITS INCOME[b]

	I	II
Marginal Propensity to Consume	0.2	0.3
Marginal Propensity to Import	0.2	...
Marginal Propensity to Invest	0.1	0.1

Country I (the U S.)				Country II			
Consumption of Domestic Goods	Consumption of Foreign Goods	Investment	Income	Consumption of Domestic Goods	Consumption of Foreign Goods	Investment	Income
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
...	10	...	10	...	10	...	10
2	12	1	15	3	12	1	16
3	13	1.5	17.5	4.8	13	1.6	19.4
3.5	13.5	1.75	18.75	5 82	13.5	1.94	21 26
3.75	13.75	1.875	19.375	6.252	13.75	2.084	22.086
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...

(a) This table has been adapted from Metzler, Journal of Political Economy, 1942, p. 401. Column 2 consists of the increase in imports due to the marginal propensity to import plus 10, because we assume that the original increase of 10 represents a continuing increase in the average propensity to consume imported goods. The difference from Metzler's table is that he shows in each period changes from the immediately preceding period, while our table shows changes of each period as compared with the initial period. In this we follow Machlup's procedure. This is evidently a difference in form only.

(b) This table can be immediately adapted to show the effects of a continuing loan of 10 by the United States (country I) to the rest of the world (country II). We would simply add 10 units to column 3, investments, instead of column 2, imports Consumption of foreign goods of country II would be precisely as it was before, consisting of the 10 units loan plus whatever were the induced imports of country I. Nothing else need be changed in the table.

no dollar shortage[2] will arise if there was none in the original position. But if there was such a shortage originally, the increased imports of the United States obviously will not make the dollar shortage any better.[3] If the rest of the world is unstable, an initial dollar shortage will become worse with rising incomes, or, if there was no dollar shortage to begin with, an increased income in the United States and the rest of the world will create such a shortage. We must conclude, therefore, that any dollar shortage which at present exists will continue, if not get worse, in the future, even with rising American incomes for as long as the conditions assumed exist.

In such a situation a free trade policy can benefit the United States not only because of the increased international division of labor and lowered cost of imports. Starting from a situation of underemployment, a less restrictive trade policy would under the assumed circumstances lead to increased income and employment at home. On the other hand, a protectionist policy would not only lower the welfare of the country through diminished specialization and the lowered efficiency which goes with it, but it will lead directly to lowered spending, lowered money and real incomes, and lowered employment. Every decrease would lead to simultaneous decreases in exports, since the consumption function of foreign countries would shift downwards simultaneously, and no completely offsetting shift from imported to domestic goods is at all likely to occur either in the United States or abroad.[4]

2. The criterion whether a dollar shortage exists or not is in practice whether the dollar rate of foreign currencies can be maintained without exchange control. If foreign countries find themselves compelled to introduce exchange control as an alternative to exchange depreciation — aside from the problem of short-term flight capital — there is a shortage of dollars.

3. This, therefore, does not contradict Kindleberger's point that "a reduction of the American tariff is not of itself an adequate solution for the world shortage of dollars ... the earnest admonitions of the world to the United States that it 'live like a creditor nation' fail to come to grips with the fundamentals of the problem." (Op. cit., p. 381.) It seems very likely that Kindleberger is correct in his forecast that "it may be doubted that increased imports would correct for long the world shortage of dollars." (Op. cit., p. 387).

4. Retaliatory tariffs and exchange controls make the downward shift of the average propensities to import a certainty, even if our conditions are not accepted. The classical position on tariffs can be reconciled with the income-creating properties attributed to tariffs by multiplier theory by the shifts in the average propensities to import.

Mr. Mosak points to yet another slightly different case in which trade restrictions will fail to give an inflationary stimulus even in the short run,

IX. Conclusions

Making the assumptions stated above we have reached the following conclusions. A simultaneous increase in imports and exports will have an expansionary effect if it is not offset by downward changes in the average propensity to consume domestic goods. Therefore, the effects of foreign trade on national income and employment are not completely described by the effects of trade balances. The best way to describe the effects of trade balances is by means of the marginal propensity to import. The best way of describing the effect of the volume of trade as distinct from the trade balance is by means of the average propensity to spend. For a discussion of the full effects of foreign trade on national income, both average and marginal propensities have to be considered, just as in ordinary multiplier discussion not only the slope of the consumption function is important but also its height and shifts. Thus the discussion presented supplements the usual foreign trade multiplier analysis.

The guess may be ventured that Harrod[5] may have had such a distinction in mind. Similarly, Haberler's conclusion seems too pessimistic when he writes: "The point is that what will happen 'when both exports and imports change simultaneously' depends on special conditions, whilst in the case of an excess of exports over imports (. . .) the primary effect is clearly stimulating. (. . .) It follows that on the level of abstraction and simplification on which the multiplier analysis is carried out traditionally . . . only an excess of exports over imports can be regarded as a stimulating

and even if foreign countries do not retaliate. "It has already been pointed out that the direct effect of the import restrictions on the rest of the world is deflationary. If their exports constitute an important fraction of their income, the secondary repercussions may aggravate the contraction to such an extent as to lead to a general world depression which spreads to the first country as well. This is particularly likely under elastic price expectations." (Jakob L. Mosak, General Equilibrium Theory in International Trade, pp. 172–173.) This case is related both to the case which Machlup discusses (Op. cit., p. 32) through expectations, and to Metzler's unstable case through the secondary repercussions. But, to repeat, our results depend neither on expectations nor on lags nor on instabilities.

It might be added that the present insistence of the State Department on freer international trade as income-creating finds thus support in both orthodox and multiplier theory.

5. R. F. Harrod, International Economics, Rev. Ed., Chs. 6, 7.

factor, whilst the parallel shift upward of both exports and imports must be assumed to be neutral."[6]

A balanced increase in trade will, under the assumptions made, change the average propensity to consume, or the average propensity to invest, or both. The expansionary effect of an autonomous increase in the average propensity to consume and/or to invest is independent of any lag between imports and exports. Nor are any instabilities involved. The effects of such a lag, or of such instabilities, are additional.

It is often said that free trade is clearly beneficial with full employment, but that with underemployment tariffs and other trade restrictions may increase the welfare of the country which imposes them. It has been shown that under condition of underemployment, and with assumptions which seem realistic and no more restrictive than the ones usually made, this is not necessarily so or even likely to be the case. It seems probable that the reasoning will have important policy implications in the postwar period, after the initial postwar boom has passed.

<div align="right">WOLFGANG F. STOLPER.</div>

SWARTHMORE COLLEGE   •

6. G. Haberler, Prosperity and Depression, 1941, p. 468. Haberler adds the following qualifying footnote: "Those who maintain that changes in exports must be put into the multiplicand irrespective of whether they are or are not accompanied by parallel changes in imports, may have a point; but they are stepping outside the multiplier analysis. There is no objection to that provided that the conditions under which the stimulating effects of the increase in export exceed the depressing effects of the increase in imports are precisely stated. This has, in fact, not been done. It is worth noting that, if the case is correctly stated, it follows that also a decrease in exports may have to be regarded as a stimulating factor if it is accompanied by a decrease in imports, because the stimulating effects felt by the industries favored by a fall in imports may be greater than the depression caused in the export industries."

I hope that I have succeeded in stating more precisely the conditions under which a balanced increase in imports and exports would be expansionary. From the last sentences of Haberler's footnote it appears that the case discussed in this article differs from the "special conditions" which he has in mind.

GUILD PRICING IN THE SERVICE TRADES

SUMMARY

The medieval guilds, 311. — Modern guild regulations in the service trades — the barbers: importance of legislative aid, 312; non-price regulations, 314; price-fixing laws, 316; enforcement of regulations, 323. — Economic aspects of the pricing of barber services: characteristics of the "product" and the market, 325; pricing without restriction on individual action, 328; pricing with concert of action but without benefit of legislation, 330; pricing with minimum price legislation but without extralegal concert of action, 331; pricing with minimum price legislation undergirding extralegal concert of action, 333. — Guild pricing in other trades, 335.

THE MEDIEVAL GUILDS

The modern quest for economic security by members of the service trades, figuratively, has turned back the clock to the later Middle Ages. The craft guild system, the dominant form of industrial organization from the thirteenth to the seventeenth centuries, with its "fair" wage for the producer, "honest" product at a "just" price for the consumer, and with the hazards of rivalry and change reduced to a minimum, provides the pattern for the organization of some of our present-day trades.[1]

In theory, each medieval craft guild was an administrative organization, quasi-official in nature,[2] charged with the management of a particular industry. The guild limited membership in

1. This summary of medieval guild organization is based upon the following works: Ashley, W. J., An Introduction to English Economic History and Theory, Vol. 1, parts 1 and 2 (4th ed., New York, 1906), Birnie, A., An Economic History of the British Isles (New York, 1940), Cunningham, W., The Growth of English Industry and Commerce (5th ed., Cambridge, 1910), Heaton, H., Economic History of Europe (New York, 1936), Lipson, E., The Economic History of England (5th ed., London, 1929), and Usher, A. P.. An Introduction to the Industrial History of England (New York, 1920).

2. The relationship between the guild and the government varied in detail, but practically all guilds sooner or later seem to have become administrative adjuncts to the Municipal Government, performing services not unlike modern commissions or bureaus. The guild administered a craft, but its regulations seem to have been subject to approval by governmental authorities. The actual degree of control in the public interest might vary with the particular arrangement and with the political power of the guild members. These men, after all, usually were leading citizens, frequently members of the town council, and at times were the most influential element in the local government. Lipson, op. cit., Vol. 1, pp. 338, 340–342.

the craft, and defined the terms of entrance.³ Moreover, the guild prescribed in detail the manner of operation of shops, and the general conditions of production;⁴ and it played a leading part in the determination of prices, particularly minimum prices, thus reducing the intensity of competition. In exercising this great power, the masters of the guild were supposed to keep in mind the welfare of the general public, although there is evidence that this aim often was lost sight of in favor of a more selfish point of view.⁵ The guild system in essence, then, was one in which a high degree of control was exercised over the product and its price as well as over the members of each trade by those engaged therein through the use of concerted activity, often doubtless to the disadvantage of the ultimate consumer.

Modern Guild Regulations in the Service Trades

Today the guild type of activity is returning. In no field of activity has this development made greater progress than in the barber trade. The smallness of the business unit, the large degree of homogeneity of those engaged in the trade, the background of public health regulation, and the severity of impact of depression on this business render the occupation particularly amenable to guild organization.

Perhaps the outstanding feature of the development of modern guilds is the manner in which their growth has been secured by legislative action. Lacking this endorsement, associations of mem-

3. Limitation of entrance developed with the growth of the apprentice system. This institution possessed certain advantages from the point of view of those engaged in the trade. An apprenticeship of many years meant sound training and at the same time resulted in controlling numbers in the trade.

4. In order to prevent poor work and fraudulent activities (which might bring the entire guild into disrepute), the wardens of each guild were empowered to inspect workshops and products, to fine violators, and to require that members mark their goods in such a manner as to make possible tracing the source of poor work. Furthermore, the guilds often prohibited night work, because of inspection difficulties and the chance of unsatisfactory work slipping by. Incidentally, the use of improved methods frequently was prohibited, and larger scale operations often were prevented by limitations upon the number of apprentices or journeymen to be employed by a single master.

5. For example, according to Ashley, op. cit., Part 2, p. 160, one Parliamentary statute revealed that: "masters, wardens and people of gilds, fraternities, and other companies corporate, dwelling in divers parts of the realm, often times . . . make among themselves many unlawful and unreasonable ordinances as well *in prices of wares* as other things for their own singular profit and to the common hurt and damage of the people."

bers of a trade might run afoul of state antitrust legislation, if they attempted to agree upon such matters as minimum prices, working hours, and the right to practice the occupation.[6] But these same associations have prevailed upon legislatures of many states to "protect the public welfare" by passing laws which accomplish much the same result. Administration of such statutes in the barber trade is carried on by boards of barber examiners, composed of members of the industry, whose functions include the application of provisions governing the training, examining, and licensing of tradesmen, the enforcement of rules respecting shop sanitation and trade practices, and the administering of minimum price regulations.

Typically these boards are made up of three members, generally practicing barbers, usually appointed by the governor of the state. In some instances, restrictions are placed upon the governor's freedom of selection. For example, the Colorado statute specifies that one member must be recommended by the unions of journeymen barbers; a second member must be an employing barber proposed by Colorado chapters of the associated master barbers of the state; the governor may make his own choice in the selection of the third barber member of the board.[7] Since the ostensible purpose of this type of regulation is the protection of the public health and welfare, it is significant that few statutes make any provision for consumer representation on the board.[8]

The modern barber laws, then, establish a relationship between industry and government akin to that obtained by the craft guilds of the later Middle Ages. If anything, the modern approach grants even greater autonomy to the trade, in that the administrative authority, once established by the legislature, thereafter is relatively free from responsibility to its creator. For example, some

6. The status of service trades with respect to state antitrust laws is summarized briefly in footnote 9, infra, p. 331.

7. Stat. Ann. (Michie, 1935), c. 19. Extensive discussions of the control of the regulating authority by members of the trades and professions is contained in Grant, J. A. C., "The Gild Returns to America," Journal of Politics, Vol. 4, pp. 303–336, 458–477 (1942). This study, while not covering the price aspect of modern guild activities, supplied much of the inspiration for the present analysis. See also Fellman, D., "A Case Study in Administrative Law — The Regulation of Barbers," 26 Wash. U. L. Q., pp. 213–242 (1941).

8. Of the states which have passed laws empowering barber boards to set minimum prices, only Wisconsin requires that a member of the board be chosen from the general public. Wisconsin Laws (1939), c. 514.

laws provide that the fees collected by the agencies for licenses and examinations constitute a fund to be used for the support of the board, thus apparently freeing it from the necessity of coming before the legislature annually for an appropriation.[9] And while the boards ordinarily must make a report to the governor each year, summarizing their activities, these reports are so general and routine that their requirement can hardly be considered a restriction on freedom of action.[1] The courts exercise practically the only check on arbitrariness of action by the boards through their function of reviewing the power granted to, and the administrative decisions made by, the board, upon prayer by an aggrieved party. Thus, barber boards have become a part of government, and as such they may compel all members of the occupation to conform to board regulation. The nature and extent of this regulation, of course, depends upon the character of the statutes which the individual board must administer.

Non-Price Regulations. The provisions of laws regulating barbering fall into three progressively restrictive classifications. The first class includes those obviously desirable measures which require barbers to observe certain elementary health and sanitary precautions. The main purpose of these provisions is not restrictive in nature; they may increase somewhat the cost of service, but they bear alike on all members of the occupation. Many of these regulations, such as that requiring a barber to use a fresh towel on each customer, originated as rulings issued by health or sanitary commissions. With the enactment of barber codes containing sanitation provisions, barber boards have acquired some authority in this area.[2] Insofar as the board has jurisdiction in this matter,

9. Among others, the statutes of Louisiana, New Mexico, North Dakota, and Oklahoma so provide. Citations for these laws are given in footnote 2, infra, p. 317.

1. The eight-page Annual Report of the California Board of Barber Examiners is devoted chiefly to a detailed statement of income and expense, by months. A very brief summary of non-financial activities, a table showing the number of barbers, barber shops, and population in each county, and a paragraph on penalty proceedings complete the Report.

2. A substantial amount of overlapping exists with respect to the authority over sanitation possessed by state or local boards of health and the barber boards. In Los Angeles, for example, the Board of Health annually inspects and licenses all barber shops within the city limits; however, the California Barber Board is empowered to do the same thing, and every year takes disciplinary action against shop owners whom it finds to be violating its rules in this matter.

its regulations may be likened to medieval provisions governing the conditions of production, inspection of work, and the trade-marking of articles.

The second class of regulations is somewhat restrictive in character. The most fundamental step in this direction is taken by those statutes, almost universal now, which require that no person may practice the barbering trade without first obtaining a license. More and more the acquisition of a license, or at least its retention, has been hedged about with restrictions. Usually an applicant for a license must possess certain qualifications: in California he is required by law to be of good moral character,[3] to be free from contagious disease,[4] to have had "adequate" training or experience,[5] and to be able to demonstrate his skill to the satisfaction of the administrative agency.[6] It is not easy to estimate how effectively these various measures limit membership in the trade, but their tendency is clear.[7]

3. In State vs. Walker, 92 P 775, 777 (Washington, 1907), a dissenting judge stated that he failed to see any necessary relationship between the public health and the morals of a barber, however determined. But most courts have given general, if somewhat vague, approval to this qualification.

4. The early boards of health usually required this safeguard, and its obvious desirability has led to its uniform acceptance by the courts. It must be reasonable in scope, however, as the court declared in Schneider vs. Duer, 184 A 914, 920 (Maryland, 1936).

5. For example, the Maryland Barber Code required the barber to be trained in: "Scientific fundamentals for barbering, hygiene, bacteriology, histology of the hair, skin, nails muscles and nerves, structure of the head, face and neck, elementary chemistry relating to sterilization and antiseptics, disease of the skin, hair, glands and nails, haircutting, shaving and arranging, dressing, coloring, bleaching and tinting of the hair." Laws (1935), c. 371. in holding this law invalid the Supreme Court of the State said, "Such requirements are obviously far beyond the need of this trade. . . . There is an apparent design, although indefensible and unreasonable, to give this simple and useful trade the characteristics and standards of a highly technical profession . . ." Schneider vs. Duer, 184 A 914, 918 (Maryland, 1936). See also Whitcomb vs. Emerson et al., 115 P(2d) 892 (California, 1941). For a more comprehensive discussion of these requirements, see Fellman, footnote 7, supra, p. 313.

6. In the case of barbers, the examination is both written and manual. The absence of absolute standards makes possible the exercise of considerable discretion by members of the barber boards. See Fellman, op. cit.

7. The California barber code was enacted in 1927, but the State Board of Barber Examiners did not report the number of licensed barbers until 1929, at which time there were 18,459. By 1945 this figure had declined to 15,434. Expressed in per capita terms, in 1929 there were about 3.3 barbers per thousand population; this ratio declined almost steadily to 1945, when it stood at 2.2 barbers per thousand population. It is, of course, possible that other

Non-price regulatory measures which attempt to reduce the rigors of competition by various means fall into this category also. Regulatory enactments forbidding barber shops to remain open after six o'clock or to operate on Sunday are examples of this sort of restriction. Such laws or ordinances operate presumably to prevent the more aggressive members of the trade from obtaining an advantage by offering greater convenience to customers. The varying attitude of the courts toward this type of regulation is exemplified in the peculiar Idaho decision which ruled that a state law governing hours of operation was illegal but that a municipal ordinance providing for the same thing was valid.[8] The attempt to prohibit price advertising by legislative enactment is similarly restrictive in purpose. Certain laws of this type even have gone so far as to forbid the barber from posting his price list in such a position that it can be seen without entering the shop. The purpose of this type of regulation evidently is to decrease the sharpness of price competition by reducing consumer access to price information. The courts have shown no sympathy toward such measures, striking them down whenever their validity has been contested.[9]

Price-Fixing Laws. More extreme than the counterpart "Fair Trade" and "Unfair Practices" legislation in the field of retail merchandise distribution are those statutes which empower state boards of barber examiners to set minimum prices for barber services by areas within the state.[1] The California statute, enacted in

factors than restrictions on entrance to the trade have contributed to this decline.

8. Pearce et al. vs. Moffatt, 92 P(2d) 146 (Idaho, 1939).

9. For example, a provision of the Ohio statute empowering the Board of Barber Examiners to suspend, revoke, or refuse to issue a license to any barber who advertised the prices of barber services in any form whatsoever was declared unconstitutional in Jones vs. Bontempo et al. 32 N.E. (2d) 17 (Ohio, 1941). See also People vs. Osborne, 59 P(2d) 1083 (California, 1936). However, limitations on price advertising may remain in effect for some time in the absence of court test. The Rhode Island barber service price fixing law, cited in footnote 2, p. 317, infra, forbids price advertising and has remained uncontested since its enactment in 1942.

1. The acts aimed at reducing the stringency of price competition between retail institutions fall short of barber pricing codes in several respects. Horizontal organization among retailers, i.e. concerted action, is forbidden rather than encouraged by the retailer laws. Moreover (1) the so-called "Unfair Practices Acts" do not set uniform minimum prices, but permit differentiation in prices as between dealers on the basis of individual variations in costs of operation; and (2) although the "Fair Trade" laws specify a uniform minimum price, their application is limited to branded products, and then only at the option of the brand-owner. In addition to the foregoing is the fact

1941, and in full operation though as yet inadequately tested in the courts, is typical in most particulars of laws of this kind in other states.² It begins with the following declaration of policy by the legislature:

> In enacting this statute the Legislature finds that unfair, unjust, destructive, demoralizing and uneconomic trade practices have been and are now being carried on in the operation of barber shops in the State of California, and that unfair competition exists between the individual barber shop owners and managers of this State to the extent that prices have been reduced by such unfair competition to the point where it is impossible for an average barber, although working regularly, to support and maintain reasonably safe and healthful barbering services to the public.
>
> Such conditions constitute a menace to the health, welfare and reasonable comfort of the inhabitants of this State, and tend to the transmission of

> As the barber business affects the health, comfort and well-being of our citizens, and of the public which patronize the barbers of the State of California, in order to promote the public welfare, health and safety, and to prevent the transmission of disease, in view of the personal touch and contacts manifested and exercised in the barber business, and the need for well-nourished, strong and healthy persons to engage in the barber business, the barber profession is hereby declared to be a business affecting the public health, public interest and public safety.*

that, unlike the merchandise field, when competition is eliminated at the retail level in the service trades, there is none left.

2. Cal. Stat. (1941), c. 317. Other states which have enacted what may be loosely referred to as the "model" barber pricing statute include: Arizona, Laws (1939), c. 38; *Arkansas, Acts (1941), Act 432; Colorado, Laws (1937), c. 105; *Florida, Laws (1941), c. 20425; *Indiana, Acts (1939), c. 108 (repealed) and Acts (1941), c. 77; Kansas, Laws (1941), c. 298; Louisiana, Acts (1936), No. 48; *Michigan, Acts (1941), No. 309; Montana, Laws (1939), c. 150; New Mexico, Laws (1937), c. 230; North Dakota, Laws (1943), c. 93; Oklahoma, Laws (1937), c. 24, Art. 2; *Oregon, Laws (1943), c. 330 (repealed), and Laws (1945), c. 198; Rhode Island, Acts (1942), c. 1253; *Tennessee, Acts (1937), c. 236. In addition, two other states, Minnesota, Laws (1937), c. 235, and *Wisconsin, Laws (1939), c. 514, have passed laws providing for the setting of prices for certain service trades, including barbering. Fundamentally these laws are very similar to the special barber statutes. Adverse court decisions have invalidated enactments in those states marked with an asterisk. It appears, then, that of the eighteen barber pricing laws passed to date, only eleven are still in effect. Three other pricing statutes, different in character although applying to the service trades, are not included in the above list. See footnote 5, infra, p. 318.

3. Section 1. The declaration changes tack at this point to a "let's get our noses into the trough also" philosophy as follows:

> "In recent years social security, a ceiling for hours and a floor for wages, collective bargaining, prohibition against unfair trade practices including sales below cost, and validation of resale price maintenance agreements, together with many other reforms have been provided for many industries

In recent years all pricing laws for the service trades have included labored introductions of this kind, which, in the quotation employed by one judge, "doth protest too much, methinks."[4] Apparently this approach has been used in an attempt, often successful, to avoid the adverse court reaction which greeted three relatively early price-fixing laws for the service trades based upon an emergency need,[5] and to bring the act within the scope of the Nebbia decision.[6]

and businesses, resulting in an increased standard of living for persons engaged in them and a consequent raising of the standards of service of those industries and businesses to the public in terms of competency and sanitation.

These reforms do not generally apply to barbers. Barber shops are usually very small establishments, very often owned and operated entirely by one man. Barbers do not sell a trade-marked commodity; they are not in interstate commerce.

In order then, that the state may do its part in protecting the public and in insuring a nondiscriminatory application of the recent increased movement to achieve social progress, this statute is enacted."

4. Board of Barber Examiners of Louisiana vs. Parker, 182 So. 485, 514 (Louisiana, 1938).

5. The laws of Alabama, Gen. Acts (1935), pp. 746–748, California, Gen. Laws (Deering), 1937, Act 8784, and Iowa, Code (1939), c. 292.1, stated that an emergency situation existed among the service trades which affected commerce and the public welfare, and declared that it was necessary to provide for "fair competition" by authorizing municipalities to pass minimum price ordinances when petitioned to do so by 60 per cent (or 65 or 70 per cent, as the case might be) of the members of a particular trade. All three of these laws were declared invalid, the courts holding that emergency does not extend the basic powers of the legislature and that these statutes constituted special group legislation quite beyond the constitutional power of the governments to enact. City of Mobile vs. Rouse, 173 So. 266 (Alabama, 1937), Ex parte Kazas, 70 P(2d) 962 (California, 1937), and Duncan vs. City of Des Moines, 268 N.W. 547 (Iowa, 1936). The first Florida statute, Laws (1935), c. 16799, although founded on the premise of public health protection, declared that it was enacted to meet an emergency and specified that its duration was to be only two years. It was invalidated in State ex rel. Fulton vs. Ives, 167 So. 394 (Florida, 1936).

6. Nebbia vs. New York, 291 U.S. 502 (1934). In this decision the court discarded the narrower traditional view of industries subject to price regulation and stated (pp. 536, 537) "The phrase 'affected with the public interest' can. . . mean no more than that the industry, for adequate reason, is subject to control for the public good. . . . If the laws passed are seen to have a reasonable relation to a proper legislative purpose [which include the protection of public health, safety, and welfare], and are neither arbitrary nor discriminatory, the requirements of due process are satisfied. . . ." See also West Coast Hotel Co. vs. Parrish et al., 300 U.S. 379 (1937) and Olsen vs. Nebraska, 313 U.S. 236 (1941).

State courts, however, have differed in their interpretation of the Supreme Court decisions. Some have reasoned that price-fixing legislation is constitutional in any business declared by the legislature to be "affected with a

The argument that prices must be fixed for barber services, in order to protect the public health and safety, has been stated in various forms by those courts which have upheld these laws, but in general it runs something like this. The barber and his equipment come into physical contact with the person of his customer; it is essential from the standpoint of public well-being, therefore, that the barber be healthy and that his equipment be sanitary. If prices for barber services are too low, it is contended, not only will the barber's health be impaired, with a resulting danger to customers, but he will be unable to afford to keep his tools in a sanitary condition. It follows, then, that prices must cover the costs of maintaining a high standard of health and sanitation. This argument blithely ignores the critically important possibility that the demand curve for barber services may be elastic above prevailing prices, which will materially affect volume of operations,[7] overlooks the fact that the state already possesses laws, assertedly enforced, providing for compulsory sanitation and inspection, and finally, does not take cognizance of the fact that disinfectants may

public interest," i.e. subject to regulation for the public welfare. The majority members of the supreme courts of five states, Florida, Louisiana, Minnesota, New Mexico, and Oklahoma, have adopted this viewpoint in approving price provisions in barber codes. Cf. McRae et al. vs. Robbins et al., 9 So. (2d) 284 (Florida, 1942), Board of Barber Examiners of Louisiana vs. Parker, 182 So. 485 (Louisiana, 1938), State vs. McMasters, 283 N.W. 767 (Minnesota, 1939), Arnold vs. Board of Barber Examiners et al., 109 P(2d) 779 (New Mexico, 1941), Herrin et al. vs. Arnold, 82 P(2d) 977 (Oklahoma, 1938). On the other hand four state courts, those of Arkansas, Indiana, Tennessee, and Wisconsin, have taken the position that the Nebbia ruling means only that price fixing statutes can be valid in those occupations or under those circumstances in which it is apparent that there is a clear and definite relationship between the fixing of prices and the protection of the public health, and they do not see this necessary connection in the barber trade. Cf. Noble et al. vs. Davis, 161 S.W. (2d) 189 (Arkansas, 1942), State Board of Barber Examiners vs. Cloud et al. 44 N.E. (2d) 972 (Indiana, 1942), State vs. Greeson et al., 124 S.W. (2d) 253. (Tennessee, 1939), and State vs. Neveau, 294 N.W. 796 (Wisconsin, 1940). In seven states barber price statutes remain untested by the courts, and in three others, California, Michigan, and Oregon, the fundamental question as to the state's power to provide for such pricing has not been decided upon in those cases which have appeared before the courts.

7. As a matter of fact, the price of shaves in Los Angeles is so high (the legal minimum is 35 cents, the typical price, 75 cents) that the number sold is negligible, although the fact that in some instances barbers are unwilling or unable to perform the service may contribute to this absence of volume. See footnote 2, infra, p. 327.

be purchased for so little that income and sanitation bear little if any relation to each other.[8]

After stating the legislative intent to protect the public health and welfare, the California law, typical of many, grants the State Board of Barber Examiners certain general powers, and then specifies the following provisions regarding minimum prices:

> The board may establish minimum price schedules for the various items of barber services for any city or county upon its own motion, or upon the filing of a petition with it, requesting a minimum price schedule for that city or county, signed by 75 per centum of the barbers in that city or county.[9]

In providing that the minimum pricing process may be invoked either upon the Board's initiative or upon petition by members of the trade, California's statute has followed a course midway between two somewhat extreme positions.[1] The most conservative of these, in the sense of delegating authority in price matters to members of the trade, specifies that the decision to impose minimum price schedules rests solely with the board on its own initiative.[2] The other, and much more popular, position provides that the law is to become operative only upon petition to the board by a given percentage (varying from 50 to 80, but usually 75) of the licensed barbers of the area.[3] Indeed, most of these laws specify that the

8. In the words of one judge, "Cleanliness and antisepsis are inexpensive. Limited amounts of forty per cent formaldehyde, two per cent carbolic acid solution and soap are the only disinfectants that need be purchased under the Board's present rules. Laundry and the labor of scrubbing seem to be the only other necessary elements of cleanliness." State Board of Barber Examiners vs. Cloud et al., 44 N.E. (2d) 972, 977 (Indiana, 1942).

9. Section 2, §6556.

1. The Kansas and second Oregon enactments are somewhat similar to the California law in this respect.

2. This provision was incorporated in the first Florida law, the second Indiana measure, and the statutes of Michigan and Wisconsin. Each of these has been declared to be unconstitutional, but on grounds other than the retention of authority by the board. See footnote 2, p. 317, supra, for citations to the laws. See also, State ex rel. Fulton vs. Ives et al., 167 So. 394 (Florida, 1936), State Board of Barber Examiners vs. Cloud et al., 44 N.E. (2d) 972 (Indiana, 1942), Klosinski vs. Michigan State Board of Examiners of Barbers, 13 N.W. (2d) 211 (Michigan, 1944), and State vs. Neveau, 294 N.W. 796 (Wisconsin, 1940).

3. For example, the laws of Arizona, *Arkansas, Colorado, *Florida (second law), *Indiana (first law), Louisiana, Minnesota, Montana, New Mexico, North Dakota, Oklahoma, *Oregon (first law), Rhode Island, and *Tennessee so specify. (The asterisk denotes that the law has been held to be invalid.) As might be expected, the initiative necessary to secure signatures on petitions for board action is provided by the leaders of barber associations. Thus, executives of the Los Angeles chapters of the Journeymen Barbers'

board may only approve or disapprove out-and-out *agreements*[4] as
to minimum prices submitted by the petitioners.[5] One may specu-
late on the possibility of attempts being made to go one step
further in limiting the powers of the board, that is, to strip the
agency of its authority to veto an agreement, leaving it no choice
but to approve all prices set by the majority members of the trade.

Under present laws, however, it is the responsibility of barber

Union and the Master Barbers' Association joined forces to prepare petitions
and send representatives to every barber in the area, in order to obtain the
acquiescence of the required 75 per cent of the membership of the trade in the
area.

4. The specific requiring of minimum price agreements is to be found in
the statutes of Arizona, Colorado, *Indiana (first law), Louisiana, Montana,
New Mexico, Oklahoma, *Oregon (first law), Rhode Island, and *Tennessee.
(The asterisk indicates that the statute has been held invalid.) The *Arkansas
enactment and the second Oregon law go almost as far in this direction; they
require that a suggested list of minimum prices accompany the necessary
petition to the board requesting the setting of minimum prices.

An analysis of these statutes reveals that all barber pricing laws now in
effect permit, indeed in most cases require that barbers engage in legalized
concerted action of some sort. There are, theoretically at least, three stages of
such concerted action: (*a*) that in which barbers petition for action by the
board, but do not specify any particular prices; (*b*) that in which barbers
petition for action by the board, with a specified set of prices mentioned; and
(*c*) that in which barbers enter into an agreement concerning minimum prices
and communicate their pleasure to the board for approval. It may be that the
only difference, if any, between (*b*) and (*c*) lies in the legally binding effect of
the agreement on signatories' prices. In one case involving an agreement type
law, for example, the Supreme Court of Florida refused to consider the peti-
tioner's contention that the law was unconstitutional, holding that it was
unnecessary to rule upon this question because the petitioner must abide by
the minimum price agreement which he had signed. Cf. Economy Cash &
Carry Cleaners, Inc. et al. vs. Florida Dry Cleaning, Dyeing, and Pressing
Board et al., 174 So. 829 (Florida, 1937).

5. It is not astonishing that some courts have considered that such a
"rubber stamp" procedure constitutes illegal delegation of authority. For
example, the Indiana Supreme Court pointed out that the minimum prices
assertedly essential to the welfare of the general public would become operative
only at the instigation of persons whose interest in the matter was purely
selfish. Hollingsworth vs. State Board of Barber Examiners, 28 N.E. (2d) 64
(Indiana, 1940). Cf. also Noble vs. Davis, 161 S.W. (2d) 189 (Arkansas, 1942)
and La Forge vs. Ellis et al., 154 P(2d) 844 (Oregon, 1945). On the other
hand, a number of courts apparently have not been disturbed by this procedure
because they have approved laws providing for it. Cf. Board of Barber
Examiners of Louisiana vs. Parker, 182 So. 485 (Louisiana, 1938) and Herrin
et al. vs. Arnold, 82 P(2d) 977 (Oklahoma, 1938). In view of the fact that
the United States Supreme Court has upheld Federal legislation containing
similar provisions, it appears that the latter courts are not in error in this
respect. Cf. Currin vs. Wallace, 306 U.S. 1 (1939), United States vs. Rock
Royal Co-Op, 307 U.S. 533 (1939), and Mulford vs. Smith, 307 U.S. 38 (1939).

boards to set or at least review the setting of minimum prices. Certain standards have been established for the guidance of the boards in performing this function. The following sections of the California statute again are typical of most other legislation on the subject:

> Upon receipt of a petition under this article, the board shall investigate and ascertain those minimum prices which will enable barbers in that city or county to furnish modern and healthful services, using modern appliances and equipment, so as to minimize the danger to public health and safety incident to such services.
>
> In establishing a minimum price schedule, the board shall consider all conditions affecting the practice of barbering in that city and county, and the relation of those conditions to the public health and safety . . . the board [also] shall consider the necessary costs incurred in that city or county in maintaining a barber shop in a clean, healthful and sanitary condition.[6]

Most unbiased observers will agree that these standards permit the board to exercise considerable discretion in administering the law. This is perhaps necessary, and the degree of latitute of interpretation is certainly no greater than that granted to the Federal Trade Commission in proscribing "unfair methods of competition . . . and unfair or deceptive acts or practices in commerce."[7]

In connection with the provisions for minimum pricing, an interesting question has arisen regarding the possibility of reducing minima established in accordance with the law, if and when such a course is indicated. Often no provision is made for such an eventuality in barber service price-fixing laws. The legislature may have had this problem in mind when the following section of the California Barber Law was drafted:

> If the board, after investigation made either upon its own initiative or upon the complaint of fifty-one per cent of the barbers in the city or county for which the minimum price schedule is established, determines that the minimum prices so established, or any of them, are insufficient properly to provide healthful services to the public and to maintain a sanitary barber shop, *or that any minimum price set creates an undue hardship on barber shop owners and operators,* the board may vary or refix the minimum price for any barber service in that city or county. (Italics ours.)[8]

6. Section 2, §6556, subsections 4, 5, and 6.

7. 38 Stat. at L. 717, 719 (1914) as amended by 52 Stat. at L. 111 (1938).

8. Section 2, §6557. In refixing minimum prices, the boards usually are required to follow approximately the same procedure as that set up for original action in this regard. Under California law, as indicated above, a change is made in the percentage of the members of the occupation who must request

Assuming that the italicized phrase gives the Board power to reduce as well as increase minimum prices, when conditions warrant such action, the California law may be said to reflect a degree of enlightenment on the part of the draftsmen not present in those responsible for the writing of the statutes of some other states. Most laws have not been framed with so much foresight; the proponents of most enactments were so intent upon insuring higher prices that they failed to realize that achievement of their goal might in the long run reduce their income. This dilemma became so acute in New Mexico, which operates under a similar law in an allied trade, that the Board did lower the minimum prices, in spite of the fact that the enactment did not specifically provide for such action. On appeal, the court justified the Board's move by a lengthy and rather strange interpretation of the statute's use of the word "insufficient."[9] In reality, the court said, this provision should be interpreted to mean that prices must be such as to provide the members of the trade with enough income to fulfill the purposes of the act. If high prices so reduced volume of business that income was impaired, the court argued, then they were "insufficient" and the Board was correct in reducing them!

Enforcement. Statutory provisions regulating the conduct of investigations and hearings by barber boards are often extensive. The California Barber Law, for example, requires that the Board's procedure in gathering evidence "shall provide for a reasonable notice to all persons likely to be affected by orders to be made by the board after such investigation, for opportunity to be heard either in person or by counsel, and for opportunity to introduce testimony in their behalf at a public hearing to be held for that purpose."[1] Omission of such safeguards in one statute led the court to hold it invalid.[2] At times irregularities in the procedures followed by some boards have demonstrated the necessity of including such provisions in these statutes.[3]

the adjustment (reduced from 75 per cent to 51 per cent). A few laws merely specify that a "representative group" may petition to have the board act.

9. Nissen et al. vs. Miller et al., 105 P(2d) 324 (New Mexico, 1940).

1. Section 2, §6555.

2. That of Florida. See Robbins et al. v. Webb's Cut Rate Drug Co., 16 So. (2d) 121 (Florida, 1943).

3. The procedure used by the Indiana Board of Barber Examiners is illustrative of such irregularities. It announced that it was about to set minimum prices for barbering services in the city of Huntington, in which there were six downtown and six outlying shops. A hearing was ordered to obtain

A final aspect of the barber laws is their treatment of violations. Most laws, including the California enactment, provide such drastic penalties for violation of minimum prices that the barber has no choice but to conform, assuming that the law is enforced. For example, violations often constitute misdemeanors subject to fines of from $25 to $300 and/or imprisonment of from ten days to one year. Typically, each day's operation of a barber shop at prices below the minima constitutes a separate violation. Furthermore, the license of an offending barber may be suspended, revoked, or renewal denied. Thus, he may lose his means of earning a livelihood. The severity of these penalties and the absence of complexity of price schedules in the barber trade simplify problems of policing, which is undertaken unofficially by trade organizations and supplemented by voluntary reporting by members of the occupation. Because of the very real possibility that barbers would not risk continuing penalties in order to contest an action by the boards in the courts, some laws make specific provision that a minimum price order may be appealed to the courts within twenty days, during which time the order may be suspended.[4]

Today, associations of barbers harp continuously on the need for the prevention of "uneconomic" practices and for "fair" prices.[5] Some groups have gone beyond mere exhortation of members. The Associated Master Barbers' group continually is engaged in an active program of "education" designed to instill in barbers the concept that their occupation is a profession whose members should receive a "worthy" return. Executives of the Association personally contact members of the trade from time to time, pointing out the disadvantages to all of "unfair trade practices," and the need for the maintenance of an "ethical" shop. The Journeymen Barbers' Union maintains more effective control over its section of the trade through the union contract with

evidence regarding proper minimum prices. But the announcements were printed and mailed by the owner of one of the downtown shops (he was not a member of the Board), and the hearing was held in secret with four of the downtown shop owners present. The Supreme Court of Indiana condemned the procedure, and incidentally struck down the law. State Board of Barber Examiners vs. Cloud et al., 44 N.E. (2d) 972 (Indiana, 1942).

4. For example, the North Dakota statute so provides (cited in footnote 2, p. 317, supra). Other state laws include provision for review, but omit any suspension of the operation of the order during the period of appeal.

5. See, for example, Master Barber Magazine, issues from February 8 to April 19, 1939, inclusive.

shopowners. The standard collective bargaining agreement between the union shop proprietor and Local No. 295 (Los Angeles), representing employee barbers, contains a clause which specifies the schedule of prices to be charged by all union shops. These prices, needless to say, are higher than the minima established under the provisions of the state law.[6] The proprietor of a one-man shop must conform to the union price list, if he wishes to obtain a "union shop" card, the display of which may be essential in strong union localities. Of course, the degree to which union activity can influence price depends largely upon the extent of unionization of barber shops in each area.[7]

Concerted action of any sort that is taken to obtain relief from the pressure of price competition may have long-run results not looked for by members of the group. This appears to be particularly true of the barber trade, because of the "product" characteristics which influence the determination of prices for its services.

ECONOMIC ASPECTS OF THE PRICING OF BARBER SERVICES

Characteristics of the "Product" and the Market. The "products" of the barbering trade, so to speak, are personal services, the most important of which are haircutting and shaving. These services

6. In Los Angeles, for example, the union shop must charge $1.00 for haircuts (75 cents for children, except on Saturdays) and 65 cents for shaves; the legal minima are 65 cents and 35 cents for haircuts and shaves, respectively. In at least one court case, the legality of a collective bargaining agreement specifying prices has been upheld. The Supreme Court of Arizona required shop owners signatory to such an agreement to abide by the price schedule, in spite of their contention that such a contract was invalid as against public policy, because of its tendency to restrain trade. Henderson vs. Ugalde et al., 147 P(2d) 490 (Arizona, 1944). But cf. Cleaners, Dyers, and Pressers Local Union, etc. vs. G. H. W. Cleaners and Dyers, Inc., 7 So. (2d) 623 (Louisiana, 1942). The condemnation of certain union-management agreements regarding product pricing by the United States Supreme Court in Allen-Bradley Co. et al. vs. Local No. 3, I.B.E.W., 325 U.S. 797 (1945), may cast some doubt on the legality of these contracts, but an important feature of the Allen-Bradley case, combination among employers in their negotiations with the union, is absent in the barber situation. Furthermore, employed barbers are paid a commission on the services performed; hence their wage depends upon the price of the service, and the court might hold that the fixing of the price was a "legitimate" union objective.

7. Union price regulations may bear heavily upon smaller, less attractive shops, and proprietors of such establishments usually oppose unionization. Thus, in Los Angeles, the majority of the "East Side" shops were found to be non-union. It is reported that in some cities, particularly Chicago, groups have emulated the medieval guildsmen in resorting to violence to ensure group action.

must be performed at the point of sale; they cannot be transported or transferred. Competition, unlike the situation in merchandise lines, occurs only at a single level, that is, no competition exists between distributors as distinct from and in addition to competition between producers. Storage of barber services is impossible; hence these services are perishable in the labor sense, and loss of patronage resulting in idle time tends to mean irrevocable loss of income to the barber.

In the purchase of barber services, convenience to the buyer is of great importance; the location of a shop convenient to the buyers' homes or places of work is essential in most instances. And since, in the absence of artificial restrictions, entry into the trade is fairly easy (the practice of barbering requiring only a limited amount of skill and training and the investment necessary to open a barber shop being small), additional shops may be expected to appear wherever barbers believe they can attract sufficient local patronage to meet the modest expenses of operation, including a living wage for the shop owner.[8] The number of customers any business unit must obtain need not be excessively large, since the relatively low unit price of the service sold is somewhat offset by the repetitive nature of the demand. ·

As a result of these factors, the typical geographical pattern of barber shop location is one of considerable dispersion, accompanied by a certain amount of local concentration in primary and secondary shopping centers. The entire metropolitan market, then, is broken up into a number of relatively small areas, each occupied at most by but a few shops. Thus, oligopolistic conditions exist within each district. These small markets are not mutually exclusive, however; typically they overlap, making possible the transmittal of the effect of price changes from one to another.

The services performed by barbers are not homogeneous. Mention already has been made of differences in shop location, size, and facilities. To these must be added variations in the

---

8. On the other hand, the convenience of the consumer requires that the barber shop be large enough to minimize waiting for service. And even more important, only the large shops ordinarily obtain enough volume to support an extensive variety of associated services, e g. manicuring and shoe shining. These latter factors exert some limiting influence on the opening of additional shops in an area, and tend to encourage the increase in size of shops already in operation.

appearance of the shop, both exterior and interior.[9] This absence of homogeneity creates "product" differentiation that profoundly affects the pricing problem. The degree to which an individual's buying habits are influenced by variations among shops depends, of course, upon the importance he attaches to the various differentiating factors, which, in turn, depends to some extent upon his income in relation to the prices charged in the particular market.[1] Stated in another way, the strength of his preference for a particular shop in terms of the price premium he will pay at that shop will be governed largely by the relative unattractiveness of available alternative offerings or opportunities. These include a reduction in the amount of the service he will consume and an attempt to obtain service at a lower cost.[2] The sum of the decisions made by all individuals in the market with respect to these alternatives determines, of course, the position and shape of the demand curves for the barber shops in the market, both of which vary with changing conditions.[3]

9. Furthermore, it should be noted that even individual barbers within a shop are significantly dissimilar in personality, skill, and willingness to defer to each customer's tastes.

1. The setting of prices in the barber trade is affected to some extent by the information available to both buyers and sellers. Mass advertising media are not used extensively; information as to prices is obtained most usually by observation of shop signs and by word of mouth. Thus consumers may not be quickly aware of price changes, and as a result may continue a pattern of buying longer than would otherwise be the case. For example, a number of persons interviewed in the Westwood-West Los Angeles area did not know that certain shops had been charging 25 cents less than the "uniform" price for haircuts for over a year. In another instance, a customer of one of the lower price shops had continued his patronage after a price increase, because the proprietor had stated that "all shops were going up to $1.00." Actually, several remained at the old figure.

2. More specifically, the customer may continue to patronize his preferred shop, but reduce his consumption of the service by allowing a greater time to elapse between visits. He may even take part in some concerted opposition to high barber prices. Or he may transfer patronage from his preferred shop to a somewhat less desirable, but also less expensive, establishment, if there is one in the same district or in another market area. Of the more remote alternatives available to the consumer, the most feasible, perhaps, is the possibility of acquiring tools and performing the service himself, a course of action already adopted by consumers in the case of shaving (cf. footnote 3, p. 333). Finally, since these items are only semi-necessities, it is conceivable that the dissatisfied consumer could adopt a style which permits him to dispense with a barber service entirely.

3. It should be emphasized that a priori, at least, the nature of consumer demand is greatly influenced by changes in income. Thus, in periods of relative prosperity or "stability," a large proportion of the public may be expected

While a substantial degree of differentiation obtains among barbers in the minds of consumers, some customers of all shops are likely to be sensitive to price changes. Therefore, an individual vendor can be expected to practice some foresight when considering a change in price. That is, he is apt to hesitate about reducing prices because of possible retaliatory action by competitors, and to show some reluctance about raising prices because of a possible loss of custom if rivals fail to follow suit. However, the oligopolistic-monopolistic competitive situation (smallness of numbers combined with "product" differentiation) is far from perfect, due to incorrect estimates by individual sellers, ease of entrance into the field, and so forth, and thus a considerable degree of price competition may prevail. Group action, of course, may increase the individual's discernment, and minimum price legislation may require him to conform to a certain extent to the group pattern. Each of these restrictive influences may produce its own effect on the price, and the effect of each must be considered separately.

Pricing Without Restriction on Individual Action. In the absence of restrictions on the actions of individual barbers, prices for their services may be expected, on a basis of abstract thinking, to reflect the oligopolistic-monopolistic competitive characteristics of the trade. Thus, a sort of equilibrium may obtain, in which price variations may exist from shop to shop, or at least from one class of shop to another class.[4] These variations measure the barbers' estimates of the divergent abilities of each shop to attract patronage. Should a proprietor overestimate his shop's drawing power, and as a result set prices on his services which are too high, the ensuing decline in volume of business should, in time, inform

to place more emphasis upon convenience and shop attractiveness than upon even a substantial saving in price. On the other hand, in periods of depression, many buyers may be expected to become sensitive to differences in price of only a few cents. It follows, then, that the ability of "product" differentiation to support price variations among shops changes considerably as consumer income rises or falls.

4. Mention already has been made of the fact that several shops in the small Sawtelle shopping district of Los Angeles charge substantially different prices. Similar situations were found to exist in many other sections of the metropolitan area. And one instance was reported by the Secretary of the local chapter of the Associated Master Barbers of America in which four shops, all located in the same block, were charging, respectively, 65 cents, 75 cents, $1.00, and $1.25 for a haircut. Such a range is not apt to obtain in "poor" times.

him of his error.[5] On the other hand, a price so disproportionately low relative to the establishment's "product" advantages as to seriously cut into the volume of competing businesses, almost inevitably would lead to retaliatory action on the part of the latter, thus tending to nullify the increase in patronage gains by the low price.[6] It is possible, of course, that in such a case the amount of business done by every shop in the market area would increase sufficiently to make the new price level more profitable, either because custom had been attracted from contiguous market areas or because total consumption of the service had been expanded at the lower price. However, if barbers in neighboring districts are likely to retaliate, or if the over-all demand curve is not such as to bring about a general increase, a price cut would serve only to reduce the net revenue of each establishment in the market area. The intelligent oligopolist, in the absence of great pressure to do otherwise, then, is apt to realize that his most profitable policy is to stay "in line," that is, to price his service in such a way that he neither loses too much custom to, nor gains too much custom from, other vendors in his market.

This assumed condition of stability in the barber trade is vitiated to some extent by the relative ease with which shops may be opened and the inability of individual barbers accurately to forecast the effects of their actions. Price reductions are apt to take place, particularly in periods of reduced consumer demand. Price cutting in such cases (which, unlike its counterpart in retail merchandising, usually originates not with the large but with the *small* vendor)[7] may then continue until some barbers, perhaps many, are forced into other occupations. The effects of price cut-

5. It appears that in certain instances operators, particularly of one-chair shops, believe that approximately the same net revenue can be obtained at two different prices, the higher of the two being offset by lower volume, i.e. that the demand curve for their services possesses unit elasticity between the two prices. In such cases the proprietor usually elects to charge the higher of the two prices.

6. Mr. H. R. Huntsman, Secretary of the Los Angeles Chapter of the Associated Master Barbers of America, vigorously echoed the idea that "one rotten apple spoils the whole barrel," i.e. in a stable situation a single price cutter forces all the shops in the market to drop their prices.

7. Although mention was made by Mr. Huntsman of certain "unethical" entrepreneurs who undertake to operate on a mass production basis, that is, to operate a large shop on a low price, large volume policy in order to obtain the amount of business necessary to cover the relatively large overhead burden (rent, guarantees to journeymen, etc.).

ting are not limited to the local market, of course. The barber located in an outlying section of a market area in which price cutting has occurred may be forced to drop his price when others in his area do so, with the result that he may attract enough patronage from an adjoining area to cause barbers in that market to drop their prices. In this way price cuts are transmitted from one area to the next. Moreover, the process may be speeded up considerably by the fact that, to a certain extent, barber shops in the central shopping district may offer competition to the shops in all, or almost all, outlying districts; therefore, as downtown shops are forced to lower prices, those in other areas will tend to be affected.

Pricing with Concert of Action but without Benefit of Legislation. The effect of concerted action on prices depends upon the degree of control exerted by the group and upon the nature of demand for the services. In general, however, one would expect concerted action to result in (1) greater promptness in raising prices, because individual sellers would not have to worry about pricing themselves out of the market; (2) higher prices than would be likely in the case of sellers acting individually; and (3) more "stickiness" in prices in times of business inactivity, resulting from both (a) an absence of promptness by sellers in reducing prices and (b) smaller reductions when cuts finally are made. To the extent that prices are higher generally, and to the extent that less opportunity exists to obtain service at prices "below the market," the consumer is adversely affected by concert of action by sellers.

Complete control of pricing by group action results, in effect, in monopoly price, assuming that those responsible for pricing policy are enlightened.[8] In order to obtain the full benefits of monopoly, however, it would have to be accompanied by some power to restrict entrance to the field. In the absence of such restriction, monopoly (or rather perhaps "sub-monopoly") prices would encourage the establishment of additional shops until the volume of business done by most would be so reduced as to eliminate the element of monopoly profit. In other words, the high price might result in a greater number of establishments, each operating at a volume substantially below capacity, but without

8. Although, because of many different sellers faced with varying market and cost conditions, a price which would maximize net income for one would be ruinous to another, hence the price charged in the market probably would differ from that which would be imposed by a single seller, assuming no price discrimination.

the power to change price and so eliminate rivals. Finally, however, the pressure resulting from increased numbers of sellers almost inevitably would break down the price structure, perhaps precipitating a price war. From the standpoint of the seller, non-legislative concerted action in price matters is not safe, because of its somewhat questionable legality in many jurisdictions,[9] and because of a probable absence of unanimity of action on the part of the sellers, particularly in depression times.

Pricing with Minimum Price Legislation but without Extra-legal Concert of Action. The introduction of a "floor" under prices for barber services by legislative action may affect the trade in various ways, depending upon the circumstances. If the "floor"

9. It is not clear to what extent concert of action by barbers may be deemed to be illegal. Under the common law, combinations or agreements which unreasonably suppress competition or restrain trade are illegal and void as against public policy (41 C.J. 101). Some jurisdictions apply this proscription only against restraints in the trade of necessities; others, however, have extended its scope to any lawful trade or business (41 C. J. 129). For example, in Peter Graf et al. vs. Master Horseshoers' Protective Assn., 15 O.D. 18 (Ohio, 1904) the court refused to enforce an agreement requiring that certain minimum prices be charged by the members of an association of horseshoers, holding that defendants not only did such work themselves but also employed others and therefore were not entitled to the exemption often granted labor organizations. See also footnote 6, p. 325, supra. At best, the common law does not actively prevent restraints of trade. State anti-trust laws also cannot be relied upon to render concerted action by barbers illegal. In the first place, some states, e.g. Delaware, Oregon, do not possess such enactments. Other states, e.g. Colorado, have been deprived of their anti-trust acts by virtue of court decisions ruling them unconstitutional. *Cline v. Frink Dairy Co. et al.*, 274 U.S. 445 (1927). Still other jurisdictions possess anti-trust measures whose provisions proscribe "combinations in restraint of trade in the production or sale of *articles* or *commodities*." (Italics ours.) In these instances courts may or may not include services within the scope of the law. In Missouri, for example, services, including insurance, were held to be commodities in State ex rel. Crow vs. Firemen's Fund Insurance Co. et al., 52 S.W. 595 (Missouri, 1899). See also Atlantic Cleaners and Dyers, Inc. et al. vs. United States, 286 U.S. 427 (1932). In Nebraska, on the other hand, laundries were excluded from the operation of such a law in Downing vs. Lewis, 76 N.W. 900 (Nebraska, 1898). The majority of state legislatures have employed fairly sweeping phrases in their anti-trust enactments, such as forbidding "combinations restricting trade and commerce," or even "restricting the free pursuit of any business." This language would seem to cover services, but in at least one instance a court ruled that laundries were not in commerce within the meaning of the law. State vs. McClellan, 98 So. 748 (Louisiana, 1923). As a result of these varying laws and decisions, then, barbers in a number of states seem to be relatively free to engage in concerted activities, although such activities may come within the scope of the Sherman Act to the extent that approval of local collective bargaining agreements containing minimum price provisions is required by the union's head office located in another state.

is relatively high, that is, if minimum prices are set above those previously charged by some members of the trade, it is very likely to drive these low-price shops out of business, because the basis upon which they could exist would then be reduced or even eliminated,[1] unless their better located or equipped competitors raised their prices proportionately, and even then the amount of service purchased may be affected.

Consequently, the introduction of price "floors" is likely to favor those barbers having the most attractive and most advantageously located shops. On the one hand, during periods of relative prosperity the minimum price schedule exerts a psychological pressure to force prices upward, many "superior" shops using the minima only as points of departure. That is, the poorest shops must charge the minimum prices, at least, while the proprietors of better shops exhibit a tendency to assume that they should, and can, obtain a differential above this price.[2] The widespread acceptance of this attitude makes it much more likely that uniformly higher prices can be established. Incidentally, prices may be higher with extralegal concerted action than with legal minima in good times.

The better shop also has the advantage over poorer shops in periods of stress, since the premium barber can cut his prices to the minima if necessary, knowing that he will draw customers from his less attractive or less conveniently located rivals, with the full assurance that the poorer competitor has been deprived of his only means of compensating for his "product" disadvantages — the right to reduce price. In such circumstances, then, the prices of

1. Precisely this situation occurred in Louisiana, giving rise to the first test of the barber minimum price law of that state. Cf. Board of Barber Examiners of Louisiana vs. Parker, 182 So. 485 (Louisiana, 1938). In this instance, the protesting barber, Parker by name, had operated profitably an outlying shop on prices of 25 cents and 15 cents, respectively, for haircuts and shaves. The Board of Barber Examiners established minimum prices for these services of 40 cents and 25 cents, respectively. Parker followed the new price schedule for two months, during which time his business so declined that he was no longer able to support his family. He contended that many of his customers could not afford to pay the higher prices, and that others no longer patronized his shop. Presumably, some barbers in the area benefited by the loss suffered by Parker. The court upheld the law.

2. For example, the operator of one small, old-fashioned shop explained that he felt he should charge a price above the minimum because there were shops in Los Angeles that were older, less clean, and less favorably located than his. Proprietors sometimes reflect a social as well as an economic attitude toward their prices.

barber services would tend to drop to the legal minima, and certain shops would be placed in a precarious position, as has been indicated, unless a reduction in the minima were obtained. Such a reduction is likely to be somewhat slow, since it must be based upon the gathering of evidence and the holding of hearings, preceded in some states by the filing of petitions; furthermore, the application of revised minima may be delayed by court action by one or a few barbers. Thus, in poor times prices are likely to be higher with legal minima than with extralegal concerted action.

Pricing with Minimum Price Legislation Undergirding Extralegal Concert of Action. A combination of extralegal concerted action superimposed upon minimum price schedules is likely to be the most satisfactory scheme from the point of view of the majority of sellers, in short-run terms at least. The introduction of minimum price legislation tends to reinforce greatly the price "stability" obtained by group activities. (In this setup, in fact, one great danger to the barbers is that they will overprice the service they sell, with some permanent loss of custom resulting.)[3] The schedule of legal minima removes the rock upon which most concerted activities founder — the action of a single member of a group, which demolishes the price structure by cutting prices in an attempt to gain volume at the expense of other members of the group. Of course, low minima may permit some drop from the prices set by extralegal concerted action; but the typical price cutter knows that should he start a price war, he may soon lose his only competitive advantage, the right to lower his prices still further. Moreover, he may not drop very far and remain within the law. Therefore, he is much more likely to remain "in line." Equally important, as pointed out above, the legal "floor" gives the leaders of group action the psychological support, and even the machinery, for "pegging" prices above the minima.[4] The result is that in Los Angeles, for

3. See footnote 2, p. 327, supra, regarding consumers' alternatives to paying high prices. The high prices now charged for shaves may be partially responsible, at least, for the fact that this service is no longer purchased in significant volume. Of 278 barbers in Los Angeles questioned on this point, 232 reported that they gave very few (less than four) shaves per day. This is in contrast with the situation reportedly existing twenty or more years ago, when barbers' time on some days was completely taken up with shaving.

4. That is, the legalized activities engaged in by group representatives in securing signatures on minimum price petitions to the State Board provide a precedent for, and impart a pseudo-legality to, subsequent activities respecting prices undertaken by the same individuals. The effectiveness of this technique

example, over 90 per cent of the shops in the area operate under a set of prices above the minimum price schedule.[5] A survey of 56 counties of California reveals the same relationship between actual and minimum prices.[6]

may be attested by the results of a meeting of members of the Los Angeles chapter of the Associated Master Barbers in 1944, in which it was declared that an increase in prices of barber services was warranted by changed conditions, i.e. augmented consumer income and increased costs of barber shop operation. According to an executive of the Association, within one week after the meeting 85 per cent of the membership, under no compulsion by law, adopted the proposed new schedule of prices.

5. A survey conducted in Los Angeles during the week of October 12, 1945, in which approximately one-sixth of the shops of the central metropolitan area were investigated, revealed the following distribution of prices:

| | Haircuts | | | Shaves | |
Prices	Number of Shops	Per Cent of Total	Prices	Number of Shops	Per Cent of Total
$1.25	1	.4	$1.00	8	2.8
1.00	185	64.4	.75	111	39.1
.85	4	1.4	.65	70	24.6
.75	74	25.8	.50	84	29.6
.65*	23	8.0	.35*	11	3.9
	287	100.0		284†	100.0

* The legal minima for Los Angeles County, until December 1, 1946, when they were raised to 75 cents and 50 cents, respectively.
† This total is smaller than that shown for shops reporting prices of haircuts, because the proprietors of three establishments refused to sell shaves.

As noted in footnote 6, p. 325, the union's standard prices for haircuts and shaves are, respectively, $1.00 and 65 cents. Of the barber shops surveyed in Los Angeles, 135 were reported to be union shops. Apparently a few of these proprietors failed to abide by the union price for shaves.

6. A questionnaire was sent to an apparently qualified observer (a newspaper editor or county official) in the county seat or largest city in each county of California (with the exception of Los Angeles County, which was investigated by personal survey, and Alpine County, in which no barber shops are operated). The respondent was asked to indicate the typical prices for haircuts and shaves in his town, and to state the range of prices. The data obtained from the returns to the questionnaire justify certain generalizations:

| | Haircuts | | | | | Shaves | | | |
| | Counties With Legal Minima | | Counties Without Legal Minima | | | Counties With Legal Minima | | Counties Without Legal Minima | |
Typical Prices Reported	No. of Counties	% of Total	No. of Counties	% of Total	Typical Prices Reported	No. of Counties	% of Total	No. of Counties	% of Total
$ .65*	1	2.4	0	0.0	$ .35†	3	7.1	0	0.0
.75	12	28.6	7	50.0	.50	16	38.1	9	64.3
.85	3	7.1	0	0.0	.65	2	4.8	0	0.0
1.00‡	26	61.9	7	50.0	.75	21	50.0	5	35.7
	42	100.0	14	100.0		21	100.0	14	100.0

* Of the 42 counties (in addition to Los Angeles County) possessing legal minimum prices in December, 1945, 19 had been assigned a minimum price for haircuts of 65 cents, 22 had 60 cents, and 1 had 55 cents.
† The legal minimum for shaves was 35 cents in all counties for which minima had been established.
‡ Since these data were collected, the typical price of haircuts in some communities has advanced to $1.25.

Barber legislation reduces the sharpness of price competition (already dulled by a condition of oligopolistic foresightedness) by reducing the opportunities for individual initiative in setting prices. In the service trades, reduction in competition at the retail level is particularly serious, because there is no competition at the manufacturing level (as in merchandise lines) upon which consumers can rely. Legislation in combination with non-legislative concerted action makes at once for a higher level of prices than would exist otherwise and for a reduction of opportunity on the part of consumers to satisfy their desire for this service by availing themselves of lower-priced alternative offerings. If legislation of the type described above is to be justified at all, it can only be when competition can be relied upon to keep prices within reasonable bounds; it certainly cannot be condoned when extralegal concerted activities by sellers are allowed to exist, so that monopolistic pricing is superimposed upon a system of legal minima. Actually, this is a regrettable step in the direction of price setting by a trade group whose self-interest, combined with "education" by leaders, tends strongly toward higher prices for the consumer.

Guild Pricing in Other Trades

If barber legislation were an isolated example of this type of regulatory enactment, these laws would not be particularly important. But the "progress" made by barbers serves as a beacon light for other service trades and professions, a light which, in fact, other groups have been quick to follow. The cosmetologists, for example, have organized in a fashion very similar to that of the barbers, and have obtained almost identical minimum price legislation in a number of states.[7] The similarity of the two occupations renders this achievement not hard to understand. Somewhat more surprising is the fact that proprietors of laundries

From the above information it is apparent that actual prices are above the minima in almost every case. Furthermore, typical prices tend to be higher in those counties having legal minima than in those which do not possess such a "floor" under prices. These facts furnish some corroboration, though not absolute proof, in support of the thesis that legal minima reinforce group action.

7. See Arizona Laws (1941), c. 75; Colorado Laws (1941), c. 113; Louisiana Acts (1940), No. 275; Montana Laws (1941), c. 80; New Mexico Laws (1937), c. 229; and North Dakota Session Laws (1943), c. 144. None of these statutes has undergone a direct test of constitutionality, but the New Mexico Board of Cosmetology, as was mentioned, was upheld in its action in

and cleaning and dyeing establishments also have been able to impress some state legislatures and courts with the urgency of the need for the protection of the public health and safety through the medium of minimum price legislation. These laws provide their industries with many of the important appurtenances of guild organization, including the establishment of state boards for their administration.[8] And if efforts in this direction receive legislative coöperation in barbering, cosmetology, laundering, and cleaning and dyeing, there is no reason to believe that it will stop at this point. Certainly, if these trades closely affect the public health, there are many more whose relationship to public health is at least equally close. Justice C. P. McKinney of the Tennessee Supreme Court, in a decision holding invalid minimum price legislation for barbering, stated,[9] "If the Act in question is valid, then the Legislature can directly, or through a board, fix the fees that physicians and dentists may charge for their services; the prices that hotels, restaurants and lunch counters may charge for food; the prices of meats, packing house and canning factory products; and so on ad infinitum until the liberty of the individual and the right to con-contract is destroyed." Justice Armstead Brown of the Florida Supreme Court, in a dissenting opinion opposing price fixing,

---

lowering the minimum prices for beauty parlor services, in spite of the absence of any provision in the law empowering it to do so. Nissen et al. vs. Miller et al., 105 P (2d) 324 (New Mexico, 1940).

8. See Colorado Laws (1937), c. 113; Deleware Laws, v. 40 (1935), c. 120; Florida Laws (1937), c. 17894; and New Mexico Laws (1941), c. 198. With the exception of Delaware, whose Supreme Court invalidated its law in Becker vs. State, 185 A 92 (Delaware, 1936), the principle of price-fixing for these industries has been upheld in court tests. However, two of the statutes were struck down for technical reasons, that of Colorado because it failed to provide for adequate hearings, in Smith Bros. Cleaners & Dyers, Inc. vs. People ex rel. Rogers, 119 P (2d) 623 (Colorado, 1941), and that of New Mexico because it was indefinite as to the pricing power of the Board, in State et al. vs. Alexander et al., 123 P (2d) 724 (New Mexico, 1942). Furthermore, the Florida legislature repealed that state's enactment, although its constitutionality repeatedly had been affirmed. Miami Laundry Co. vs. Florida Dry Cleaning & Laundry Board, 183 So. 759 (Florida, 1938), Florida Dry Cleaning and Laundry Board vs. Everglades Laundry, Inc. et al., 188 So. 380 (Florida, 1939), Robinson vs. Florida Dry Cleaning & Laundry Board, 194 So. 269 (Florida, 1940), and Publix Cleaners, Inc. vs. Florida Dry Cleaning & Laundry Board, 32 F. Supp. 31 (1940). Cf. Florida Laws (1943), c. 21666. It should also be mentioned that at least one early law of the N.R.A. code type (New Jersey Laws (1935), c. 281) was declared unconstitutional. Kent Stores of New Jersey vs. Wilentz, 14 F. Supp. 1 (1936).

9. State vs. Greeson, et al., 124 S.W. (2d) 253, 258 (Tennessee, 1939).

expressed a similar attitude, declaring that "if the state . . .can fix the prices to be charged by those engaged in the laundry and dry cleaning business, it likewise has the power to fix and regulate the prices to be charged by clothing stores, tailoring establishments, drug stores, grocery stores, 'the butchers and bakers and candlestick makers' . . "[1]

It is interesting to note that the Iowa emergency enactment of 1935 by its terms applied not only to barbering, cosmetology, and cleaning and dyeing, but also to medicine, surgery, osteopathy, podiatry, chiropractic, nursing dentistry, optometry, pharmacy, and embalming.[2] The present Minnesota law also has a somewhat general provision — it covers all personal service trades (excluding the healing arts) which desire to invoke its provisions.[3] Thus the groundwork for price-fixing legislation has been laid in many fields. The list of trades and professions already regulated by administrative agencies of their own choosing (but so far in nonprice matters) is growing longer. Accountants, architects, doctors, lawyers, as well as many others, have constructed, with legislative help, an organizational framework reminiscent of that of the medieval guilds.[4]

It is not only a question of how far the legislatures will attempt to go in guild pricing arrangements, but how far the courts will allow them to proceed. The United States Supreme Court as yet has not been called upon to render a decision with respect to the constitutionality of guild pricing, either directly or by denial of certiorari. Lacking an authoritative ruling in the matter, the state supreme courts have arrived at contradictory conclusions. At the present moment the highest tribunals of six states have approved the principle that the prices of service trades may be regulated; a contrary position has been taken by the supreme courts of five states.[5] Should the courts capitulate in the face of the growing

1. Miami Laundry Co. et al. vs. Florida Dry Cleaning & Laundry Board, et al., 183 So. 759, 766 (Florida, 1938).
2. Cf. footnote 5, p. 318, supra.
3. Cited in footnote 2, p. 317, supra.
4. Activities by professional groups aimed at influencing prices are not unknown. Thus, a combination of physicians to fix fees was upheld in Iowa, the court holding that medical services were a form of skilled labor not within the application of the State's anti-trust law. Rohlf vs. Kasemeier, et al., 118 N.W. 276 (Iowa, 1908).
5. The affirmative position has been taken in Colorado, Florida, Louisiana, Minnesota, New Mexico, and Oklahoma. Such pricing power has

array of price enactments, one wonders how long it will be before
guild pricing becomes the rule instead of the exception in the service
trades and professions. Indeed, will it not be apparent shortly that
the guild pricing technique is far superior to the "Unfair Practice"
and even the "Fair Trade" schemes used in merchandise fields,
from the standpoint of price "stability," and will not attempts be
made to extend its application to those areas? If so, the prospect
for the consumer, and in fact for our type of society, is indeed
bleak.

<div align="right">

WILLIAM F. BROWN.
RALPH CASSADY, JR.

</div>

NORTHWESTERN UNIVERSITY
UNIVERSITY OF CALIFORNIA, LOS ANGELES

been denounced in Arkansas, Delaware, Indiana, Tennessee, and Wisconsin.
Seven states, Arizona, California, Kansas, Montana, North Dakota, Oregon,
and Rhode Island, now possess service trade pricing laws which may give their
courts an opportunity to indicate their attitude respecting this question. This
summary does not include the early N.R.A. type laws, nor does it consider
those states whose laws have been invalidated for technical reasons, e.g.
Michigan and Oregon (first measure.)

RECENT PUBLICATIONS

ARCHI, GIAN GUALBER. Problemi in Tema di Falso nel Diritto Romano. Pavia: Facolta di Giurisprudenza. 1941. pp. 188. [Studi nelle Scienze Giuridiche e Sociali.]

ATKINS, WILLARD E.; EDWARDS, GEORGE W.; MOULTON, HAROLD G. The Regulation of the Security Markets. Washington, D. C.: The Brookings Institution. 1946. pp. 126.

BANQUE DES REGLEMENTS INTERNATIONAUX. Reglementation du Commerce de l'Or et des Devises en France. Basle, Switzerland: The Bank. 1946. pp. 200.

BAYKOV, ALEXANDER M. Soviet Foreign Trade. Princeton, New Jersey: Princeton University Press. 1946. pp. 100. $2.00.

BONNET, GEORGES. Défense de la Paix de Washington au Quai D'orsay. Geneva, Switzerland: Les Editions du Cheval Aile. 1946. pp. 390.

BROWALDH, ERNFRID. The State and the Private Banking System. Stockholm, Sweden: Svenska Handelsbanken. 1946. pp. 16. [Supplement to Svenska Handelsbanken's Index, September, 1946.]

BROWN, MILTON P. Operating Results of Department and Specialty Stores in 1945. Boston: Harvard Graduate School of Business Administration. 1946. pp. 53. $2.50.

BURNHAM, ELIZABETH A. Expenses and Profits of Limited Price Variety Stores in 1945, Chains and Independents. Boston. Harvard Graduate School of Business Administration. 1946. pp. 42 $1.50.

BUREAU OF THE CENSUS. City Finances, 1944. Washington, D. C.: United States Department of Commerce. 1946. pp. 255. 70 cents.

CANADA, DEPARTMENT OF RECONSTRUCTION AND SUPPLY. Manpower and Material Requirements for a Housing Program in Canada. Ottawa, Canada: Edmond Cloutier, C. M. G., 1946. pp. 137. 50 cents. [Published by authority of the Minister of Reconstruction and Supply.]

CARNEGIE ENDOWMENT FOR INTERNATIONAL PEACE. Year Book, 1946. Washington 6, D. C.: The Endowment. 1946. pp. 172.

CASSELMAN, P. H. The Relation between Farm Management Factors, Milk Cost, and Operator Labour Earnings. Ontario, Canada: New Era Publications. 1946. pp. 86. $1.00.

CIVILIAN PRODUCTION ADMINISTRATION, BUREAU OF DEMOBILIZATION, Labor Policies of the National Defense Advisory Commission and the Office of Production Management, May 1940 to April 1942. Washington. D. C.: War Production Board. 1946. pp. 245. [Historical Reports on War Administration, Special Study No. 23.]

——. Minutes of the Advisory Commission to the Council of National Defense, Documentary Publication No. 1. Washington, D. C.: War Production Board. 1946. pp. 181. [Historical Reports on War Administration.]

——. Minutes of the Council of the Office of Production Management. Washington, D. C.: War Production Board. 1946. pp. 102. [Historical Reports on War Administration.]

——. Minutes of the Planning Committee of the War Production Board. Washington D. C.: War Production Board. 1946. pp. 180. [Historical Reports on War Administration.]

——. Minutes of the Supply Priorities and Allocation Board. Washington, D. C.: War Production Board. 1946. pp. 75. [Historical Reports on War Administration.]

——. Minutes of the War Production Board. Washington, D. C.: War Production Board. 1946. pp. 457. [Historical Reports on War Administration.]

CLEETON, GLEN U., MASON, CHARLES W., Executive Ability, Its Discovery and Development. Yellow Springs, Ohio: The Antioch Press. 1946. pp. 540.

CLOVER, VERNON T. Trego County Real Estate Assessment Plan. Topeka, Kansas: Fort Hays Kansas State College. 1946. pp. 32. [Fort Hays Kansas State College Studies.]

CONDLIFFE, J. B. Proposals for Consideration by an International Conference on Trade and Employment. Washington, D. C.: American Enterprise Association. 1946. pp. 32. 50 cents.

CRAMER, HARALD. Mathematical Methods of Statistics. Princeton, New Jersey. Princeton University Press. 1946. pp. 575. $6.00.

CUSHMAN, ROBERT E. Keep our Press Free! New York: Public Affairs Committee, Inc. 1946. pp. 32. 10 cents. [Public Affairs Committee Pamphlet No. 123.]

DAMALAS, B. V. Monnaie et Conjoncture. Paris, France: Presses Universitaires de France. 1946. pp. 392. 300 francs.

DENTI, VITTORIO. Contributo allo Studio del Giudizio di Equità nel Processo Civile. Pavia: Facoltà di Giurisprudenza. 1944. pp. 258. [Studi nelle Scienze Giuridiche e Sociali.]

EVERETT, JOHN RUTHERFORD. Religion in Economics. New York: King's Crown Press. 1946. pp. 160. $2.50.

FEIS, HERBERT. Seen from E.A., Three International Episodes. New York, New York: Alfred A. Knopf, Inc. 1946. pp. 312. $2.75.

FRIEDRICH, CARL J. Constitutional Government and Democracy. Boston, Mass.: Ginn and Company. 1946. pp. 695. $4.00.

FROMAN, LEWIS A., Principles of Economics. Chicago, Illinois: Richard D. Irwin, Inc. 1946. pp. 1009. $5.00.

FURLAN, L. V. Das Harmoniegesetz der Statistik. Basel (Switzerland): Verlag für Recht und Gesellschaft. 1946. pp. 504.

GALBRAITH, JOHN KENNETH. Recovery in Europe. Washington, D. C.: National Planning Association. 1946. pp. 39. 25 cents. [Planning pamphlet No. 53.]

GRUCHY, CLARE DE. Creative Old Age. San Francisco, California: Old Age Counselling Center. 1946. pp. 143. $2.75.

HALM, GEORGE N. Monetary Theory. Philadelphia 5, Pennsylvania: The Blakiston Company. 1946. pp. 491. $3.50.

KUTHE, H. C. Prices and Business in 1947. New York: Business Course, 1946. pp. 271. $3.00.

KUZNETS, SIMON. National Income, A Summary of Findings. New York: National Bureau of Economic Research, Inc. 1946. pp. 144. $1.50.

LAKDAWALA, D. T. Justice in Taxation in India, with Special Reference to British Gujarat. Bombay, India: The Popular Book Depot. 1946. pp. 312. 12-8 rupees.

LERNER, ABBA P., GRAHAM, FRANK D. Planning and Paying for Full Employment. Princeton, New Jersey: Princeton University Press. 1946. pp. 222. $3.00.

LESTER, RICHARD A., AND ROBIE, EDWARD A. Wages under National and Regional Collective Bargaining. Princeton, New Jersey: Princeton University, Industrial Relations. 1946. pp. 103. Section $1.50.

LHOMME, JEAN. Utilisation, Gaspillage, Prodigalité. Paris, France: Librairie de Medicis. 1946. pp. 112.

LIU, TA-CHUNG. China's National Income, 1931-36. Washington, D. C.: The Brookings Institution. 1946. pp. 91. $1.00.

MARTIN, JAMES W. Kentucky City Finances. Lexington, Kentucky: Bureau of Business Research. 1946. pp. 275.

MERIAM, LEWIS. Relief and Social Security. Washington D. C.: The Brookings Institution. 1946. pp. 42. 50 cents.

MILLSPAUGH, ARTHUR C. Americans in Persia. Washington, D. C.: The Brookings Institution. 1946. pp. 293. $3.00.

MULLER, CHARLOTTE F. Light Metals Monopoly. New York, New York: Columbia University Press. 1946. pp. 279. $3.00.

NEW YORK, COMMITTEE ON INDUSTRIAL AND LABOR CONDITIONS. Report. Albany, New York: State of New York. 1946., pp. 210. [Legislative Document (1946) No. 51.]

NOGARO, BERTRAND. Le Developpement de La Pensée Économique. Paris, France: Librairie Générale de Droit et de Jurisprudence. 1944. pp. 345.

———. Principes de Théorie Économique. Paris, France: Librairie Générale de Droit et de Jurisprudence. 1943. pp. 308.

OWENS, RICHARD NORMAN. Business Organization and Combination. New York: Prentice-Hall. 1946. pp. 567.

PAPI, GIUSEPPE UGO. Introduction to Post-War Reconstruction Programmes. Rome: International Institute of Agriculture. 1946. pp. 265.

PETERS, CLARENCE A. International Trade: Coöperative or Competitive? New York: The H. W. Wilson Company. 1946. pp. 306. $1.25.

ROSE, H. WICKLIFFE. Tariff in International Relations. New York: The American Tariff League. 1946. pp. 16. [Publication No. 118.]

SCOVILLE, JOHN W. Labor Monopolies — or Freedom. New York City. Committee for Constitutional Government. 1946. pp. 164. $1.00.

SMITHSONIAN INSTITUTE. Annual Report of the Board of Regents. Washington, D. C.· The Institute. 1946. pp. 484. $1.75.

SOUTHARD, FRANK A., JR. The Finances of European Liberation. New York: Columbia University Press. 1946. pp. 206. $3.00.

SPAHR, WALTER E. It's Your Money. New York: Economists' National Committee on Monetary Policy. 1946. pp. 36.

STOCKING, GEORGE W., AND WATKINS, MYRON W. Cartels in Action. New York, New York: The Twentieth Century Fund. 1946. pp. 533. $4.00.

UNIVERSITY OF MELBOURNE, DEPARTMENT OF COMMERCE. Problems of Distribution in Australia. Melbourne, Australia: Melbourne University Press. 1946. pp. 40. $2.00.

UNIVERSITY OF MICHIGAN, BUREAU OF GOVERNMENT. Reorganization of Michigan's County Government. Ann Arbor, Michigan: University of Michigan Press. 1946. pp. 24. [Michigan Pamphlets No. 19.]

VICKERS, ANTONY. A Better Britain. London, England: The Broadway Press, Ltd. 1946.

VOLPE, PAUL A. The International Financial and Banking Crisis, 1931-1933. Washington, D. C.: Catholic University of America Press. 1945. pp. 123.

WHITAKER, ARTHUR P. Inter-American Affairs, 1945. New York: Columbia University Press. 1946. pp. 328. $3.75. [Annual Survey, No. 5.]

WU, Y. L. China's Economic Policy — Planning or Free Enterprise? New York, New York: Sino-International Economic Research Center. 1946. pp. 60. [Sino-International Economic Pamphlets: No. 4.]

WYCKOFF, VIOLA. The Public Works Wage Rate and Some of Its Economic effects. New York, New York: Columbia University Press. 1946. pp. 313. $3.50.

THE FALLACIES OF LORD KEYNES' GENERAL THEORY

SUMMARY

The impact of Keynes' theory, 343. — I. Summary of the Keynesian theory, 344. — Criticism, 346. — II. The effects of the demand for cash balances: a metallic regime, 347; foreign trade, 350; an inconvertible money regime, 351. — III. The origins of the Keynesian error, 354. — The lessons of experience, 357. — IV. An imperfect philosophy of unspecified rigidity, 358. — V. The political consequences of the General Theory, 362. — War-time experience, 366. — The outlook, 366.

Lord Keynes' theory, as expounded in his *General Theory of Employment, Interest, and Money,* dominates the economic thought of our time. Its author does not hesitate to declare that it demonstrates the futility of the classical theory and is destined to replace it:

I shall argue that the postulates of the classical theory are applicable to a special case only and not to the general case, the situation which it assumes being a limiting point of the possible positions of equilibrium. Moreover, the characteristics of the special case assumed by the classical theory happen not to be those of the economic society in which we actually live, with the result that its teaching is misleading and disastrous if we attempt to apply it to the facts of experience.

But the new theory has not merely a philosophical significance. It leads to rules of action, notably in the struggle against the chief malady of modern society — chronic unemployment. Indeed, it is this aspect of it — the doctrine of "full employment" — that has been most influential. Explaining the evil and providing the means of curing it, it has brought great comfort to the world.

As a remedy for unemployment, it quickly expanded beyond economic science to become an instrument of government. It has led to the publication of white papers in England and Canada and to a proposed law in the United States, the Murray Full-

Employment Bill, which undertake to bind governments to its prescriptions. The new French constitution obliges the government to present each year "a national economic plan designed to provide full employment of labor and the rational utilization of material resources." The Economic Committee of the United Nations is called "Committee on Economic Questions and Employment." Finally, the International Conference which is to deal with the problem of international trade and whose first session was held in London in October–November, 1946, is the Conference on Commerce and Employment.

The Keynesian philosophy is unquestionably the basis of a world policy today; and if the spectre of "under-employment" appears again in the world tomorrow, as is probable, it will be the universal recourse of peoples and governments. If it is true, it will be the salvation of the world; if it is false, it may lead to catastrophe by turning the world to ineffective remedies which may make the evil much worse.

For all those concerned with the future of human society there are, therefore, no questions more important at the present time than those raised by I ord Keynes' theory, and no duty more pressing than that of passing judgment on the value of the explanations which it offers and the efficacy of the remedies which it suggests. This is the task which I am undertaking here.

In formulating the criticisms which seem to me to apply to the Keynesian theory, it is a source of great regret that I must do so after the author's death. Fortunately, however, his supporters are so numerous, so active, and so powerful, that my scruples on this point are somewhat relieved. Moreover, I have already had the honor of a polemic with Lord Keynes. Far from avoiding discussion, he opened the columns of the Economic Journal to me for an article entitled, "The Ideas of Mr. Keynes on the Transfer Problem."[1]

I. The Keynesian Theory

To avoid any possibility of misrepresenting the General Theory, I quote the résumé of its doctrine given in the work itself:

The outline of our theory can be expressed as follows. When employment increases, aggregate real income is increased. The psychology of the com-

1. Economic Journal, September, 1929, Revue d'Economie Politique, July–August, 1929

munity is such that when aggregate real income is increased aggregate consumption is increased, but not by so much as income. Hence employers would make a loss if the whole of the increased employment were to be devoted to satisfying the increased demand for immediate consumption. Thus, to justify any given amount of employment there must be an amount of current investment sufficient to absorb the excess of total output over what the community chooses to consume when employment is at the given level. For unless there is this amount of investment, the receipts of the entrepreneurs will be less than is required to induce them to offer the given amount of employment. It follows, therefore, that, given what we shall call the community's propensity to consume, the equilibrium level of employment, i.e., the level at which there is no inducement to employers as a whole either to expand or to contract employment, will depend on the amount of current investment.

Thus, given the propensity to consume and the rate of new investment, there will be only one level of employment consistent with equilibrium. But there is no reason in general for expecting it to be *equal* to full employment. . . . *the economic system may find itself in stable equilibrium with N at a level below full employment.* (Pages 27–30. Italics mine.)

Such is the fundamental basis of the whole Keynesian system, the explanation of "the remarkable inability of the classical theory to serve for scientific prediction," and the demonstration of the baselessness of the "famous optimism of the traditional theory . . . founded on failure to recognize the obstacle to prosperity which may be raised by lack of effective demand."

It is a question, then, of a revolution in economic theory and a profound modification of the rules of action suggested by it. The classical theory holds that no permanent equilibrium can exist as long as there is unemployment. The Keynesian theory, on the contrary, claims that a society can continue indefinitely with large numbers of unemployed, and on this basis offers itself as an explanation of this new phenomenon in the world — chronic unemployment.

The whole Keynesian analysis is based entirely on a psychological hypothesis, the producers' insufficient propensity to consume. On this hypothesis the increase of income which might be produced by an increase of employment would not increase the demand for consumers' goods in the same proportion.

With income not giving rise to demand for consumers' goods, and in the absence of government initiatives stimulating investment expenditures of the same amount, the increment of production resulting from the increase in employment could not find a market. Lacking a market, the corresponding production would cease, and

with it would disappear the increment of income which it might have engendered. Thus would be established, by the simultaneous limitation of production and of the income making possible its acquisition, the state of under-employment equilibrium whose explanation is given as the great discovery and the essential significance of the Keynesian theory.

There is one element in this explanation which will surprise all those familiar with the analyses of the classical economics: the idea that an economic society in which an unabsorbed offer of labor exists at all times can be, without expressly assuming any fixing of prices, a state of equilibrium.

But if this is the case, it is because it is impossible in the Keynesian hypothesis that this offer of labor should be accepted, because it does not give rise in any direction whatever to any demand capable of absorbing it. Unemployment, then, is the only solution offered to the workers from which it emanates. For anyone wishing to judge the General Theory, therefore, the question is, is it possible that an offer actually appearing in the market should not give rise to any demand of the same volume? If so, Keynes' theory can explain equilibrium with under-employment, therefore chronic unemployment, and provide the means to deal with it. If not, the explanation which it offers needs to be reconsidered.

For Lord Keynes, the steps in the reasoning appear to be as follows. As a result of their insufficient propensity to consume, the workers able to take advantage of an increase in employment are not disposed to increase their expenditures on consumption in proportion to the additional income which they could obtain. Moreover, since they have no propensity to invest, they will therefore demand nothing for all the increment of resources which they do not devote to additional expenditures. I maintain that this analysis involves a serious error.

If there is really under-employment, it is not that certain workers *can* do more work, but that under the conditions offered by the market they *wish* to do more work. If they actually offer an increment of labor on the market, and if they do not intend to divert to consumption expenditure or investment the whole of the increment of income which an increment of labor makes possible, it is because they intend to increase their cash holdings by an amount equal to the increment of income which they do not spend. In proportion as they offer labor without demanding consumers'

goods or investment goods, they are, and must be, demanders of money. This is a fundamental conclusion whose necessity must be carefully understood; for we shall see that if it is admitted, it upsets the whole Keynesian construction.

If there is under-employment, it means that laborers desire to do more work. If they offer labor on the market, it is because they desire to obtain an increment of remuneration; and if they do not wish to devote their increment of resources to an increase of their expenditures on consumption or investment, it is because they intend to increase the amount of money which they keep on hand. If this were not so, their offer of labor would be purely platonic. There might be a possibility of more work, but there would be no desire for it, and there would not be under-employment.

This being so, I maintain that the demand for additional cash holdings is equivalent in its economic effects to demand for consumption goods or investment goods and, consequently, that it is able to provide a market for the labor forces offered, on the same conditions as the demand for such goods. To show this, I shall be obliged to study in detail the effect of the demand for money. This will be the purpose of the following section. It may perhaps seem out of proportion with the minor practical importance of the case with which it deals. There is no doubt that the increase of individual cash holdings could never amount to more than a limited sum, and that as soon as individuals have reached the limit of the holdings which they wish to have, they will divert any increment of resources to increasing their demand for consumers' goods or investment goods. But since the hypothesis of the non-employment of this increment of resources in response to a corresponding demand is the very center of the Keynesian argument, it is indispensable, in order to judge the latter, to study the former with care.

II. The Effects of the Demand for Cash Balances

I maintain that Lord Keynes is mistaken in claiming that incomes which do not give rise to a demand for consumption goods or investment goods, that is, which give rise to a demand for additional cash balances, will be permanently lost to the mass of incomes required for the absorption of the production associated with them and consequently will create a permanent under-employment equilibrium. To show this simply, I shall first assume

a regime where money is entirely metallic. Following that, I shall consider the general case.

If a worker enjoying an increase of employment increases his cash holdings, all other conditions, including the amount of cash holdings desired by the other members of the society, remaining unchanged, the increase in cash holdings realized by the owners of the incomes increased but not spent will necessarily have as a consequence a decrease in the cash holdings of other members of the society below the level of the holdings which they desire to maintain. To restore their cash holdings to the level desired, the latter will have no recourse but to offer without demanding. This will tend to bring about a fall in the whole system of prices.[2]

One price, however, remains stable amidst all these falling prices: the price of gold, automatically maintained at the legal parity by the purchases of the coinage authority. Hence the fall in the system of prices tends to bring about the transfer of productive resources from the products whose prices have fallen to the product whose price has not changed, a diminution in the production of the former and an increase in the production of gold. But the Bank of Issue buys all of the yellow metal offered and not demanded, and consequently supplies, by monetizing the increased production of metal, the additional cash holdings desired.

Since the fall of prices and the consequent transfer of productive resources continue as long as the cause which produced them persists — that is, the insufficiency of actual cash holdings relatively to those desired — this double movement cannot but result in bringing the former to the level of the latter by increasing the quantity of monetized metal and at the same time establishing between the price of gold, stabilized at the legal parity, and the other prices in the market the relations which formerly obtained.

Thus, the demand for additional cash holdings will have had the effect of diverting the labor forces offered in an increase in employment from the production of consumers' goods or investment goods which would not have been wanted to the production of metal destined for monetization, and consequently providing the increases in cash holdings desired.

It is therefore impossible to accept Lord Keynes' conclusion that, in the case assumed, the insufficiency of demand for con-

2. I have analyzed in detail, in Chapter 4 of my L'Ordre Social, the mechanism by which this fall is brought about.

sumers' goods or investment goods constitutes an obstacle to the increase of employment. If there is really an offer of an increment of employment on the market, and if only increases in cash holdings are desired by the persons for whom the increase of employment will provide an increase of income, the labor forces offered will find themselves spontaneously but inevitably directed by the force of the price mechanism alone towards the production of the additional cash holdings desired. Thus, the increment of production associated with an increase of employment will not have lacked a market, since it will have taken the form in which the owners of the additional incomes wished to absorb it.

It is therefore not true that the limitation of the propensity to consume, if it is not compensated by investment expenditures of an appropriate amount, is the cause of a limitation of employment. It is still less true that it leads to an equilibrium with underemployment, since the forces spontaneously brought into being by every increase in labor offered tend to adapt the economic structure to the utilization which the newly employed workers wish to make of their additional income. An economic state in process of adaptation, whatever it is, cannot be a state of equilibrium. A theory which neglects the influences tending to produce these adaptations cannot be a general theory, still less a true theory.

The Keynesian faithful will, it is true, object that the preceding analysis is purely theoretical. They will point out, first of all, that it is solely by movements of prices that the adaptation required for the absorption of an increment of production tends to be stimulated, and that in the absence of these movements or in the absence of action by price movements on the structure of the productive system, no increase of employment could be expected. Therefore, in such a case, one would, in fact, be in a state of underemployment equilibrium.

This is true, but it is no less true that, in fact, in most of the economic systems which existed before the war, spontaneous movements of prices were able to develop, and that they effectively brought about the allocation of the factors of production. The considerable variations in the rate of gold production between periods of boom and periods of depression clearly showed the sensitiveness of the productive apparatus to price movements.

I shall consider in a later section the effects of price stabiliza-

tion measures and the immobilization of the factors of production. But in no part of the General Theory are stabilization of prices and immobilization of the factors of production expressly indicated as fundamental conditions of under-employment equilibrium. If they were the fundamental conditions, it would have been indispensable that this be pointed out, for among the possible remedies it would have been necessary to count, alongside the interventions suggested by Lord Keynes, the suppression of the causes of economic rigidity. Even if this had been pointed out, however, a theory based upon such special hypotheses could not have been considered a "general theory."

In any case, even in economies not very sensitive to the forces which tend to upset economic equilibria, these forces, as long as prices are not strictly stabilized, exist, and make it impossible to consider an economic structure subject to influences which tend to modify it a state of equilibrium.

It may be noted, however, that the preceding reasoning holds only so far as workable mines of gold exist in the society under consideration. However, the absence of accessible deposits only modifies the form of the regulatory apparatus; it does not destroy it, and it eliminates none of its consequences. The fall of prices brought about by the state of under-employment, if not checked by the absorption of the under-employed into the industries producing the yellow metal, tends to divert them to the production of goods capable of being marketed abroad.[3] In this way it tends to bring about a favorable balance of payments for the country under consideration. It gives rise, as in the preceding case, to additional offers of metal on the market and consequently to additional monetizations. These latter furnish the additional cash holdings desired by the newly employed workers who do not apply their increments of income to consumption goods or investment goods.

Thus, in this case also, the fact that the workers available for an increment of employment are not disposed to devote more than a fraction of their increments of income to demand for consumers' goods or investment goods does not create a lack of markets for the increments of production which these workers can supply. It merely diverts a part of the additional production to foreign markets, where it will procure, by way of exchange, the increments

3. This mechanism, too, is analyzed in detail in Volume 1 of my L'Ordre Social (p. 383).

of metal which provide the additional cash holdings desired by the newly employed workers. In this case, again, the additional production will have been subjected to forces tending to provide it with a market. As long as prices and factors of production have not been stabilized, no state of under-employment equilibrium can exist.

The preceding analysis applies, it is true, only to a special case — that of a society using metallic money only. This leaves us with the general case of a society using inconvertible money or money which can be obtained both by the monetization of metal and the discount of commercial paper.

As in the preceding case, the non-utilization of a part of the increment of income arising from the increase of employment will lead the beneficiaries of the increments of income not utilized to increase their cash holdings. As a result, all other conditions remaining the same, the cash holdings of certain members of the society under consideration will prove to be less than they desire to hold. To bring their cash holdings back to the level they desire, they will have no other solution except to offer without demanding. It is the existence of these uncompensated offers which sets in motion a regulatory mechanism analogous in principle, if not in form, to that revealed by our study of a purely metallic regime.

The increment of offers may react either upon wealth in the strict sense or upon credit instruments. In the first case, it leads to a fall in prices; in the second, to an increase of money rates. If it affects wealth in the strict sense and credit instruments in the same proportion in which these enter into total offerings, the excess of offers resulting from the non-utilization of an increment of income will produce a fall in prices and a rise in money rates simultaneously. This preliminary statement shows the close relation which must exist between the two opposite movements. I shall next show — and this is essential for the argument — that they are inseparably bound together.

If the offer without demand reacts solely upon wealth in the strict sense, it affects cash markets to the exclusion of credit markets, since its object is the procurement of immediate increments of cash holdings. It therefore brings about a fall in cash prices. The fall in cash prices leads speculators to buy for cash with a view to sale on credit, obtaining by way of discount of the commercial paper derived from the second transaction the re-

sources required to settle the first. The increase of the demand for discount brings about a rise in rates on the money market, a rise which does not come to an end until the general level of prices stops falling.[4] Conversely, every increase in money rates leads speculators, other things remaining the same, to sell for cash with the intention of buying back on credit, investing in the market the funds derived from the first transaction until the settlement of the second. It therefore leads to a fall in the general level of cash prices.

The preceding analysis shows that the excess of offers resulting from the existence of non-utilized incomes gives rise in all cases, and simultaneously, to a fall in the general level of prices and a rise in money rates.

I know that the statement that such a relation exists will surprise certain readers who know that periods of boom, that is, periods of rising prices, are periods of high interest rates. However, the rise of money rates which has usually accompanied periods of boom in the past was caused by the increases in the discount rate decreed by the monetary authorities, almost always as a result of fears inspired by the decrease of their metallic reserves. In fact, in every country of the world, the periods of rising prices resulting from the budget deficits of recent years have been periods of very low money rates. Be it noted that in the absence of the relation stated above the functioning of an inconvertible monetary system would be simply inconceivable, since the need for cash holdings could not lead to the issue of new money. Moreover, and this seems to me the essential argument, the possibility that every excess of offers may react as well upon credit instruments as upon wealth in the strict sense suffices to make of this statement, which at first seems paradoxical, a truth of common sense.

If, in the light of this proposition, we follow the unfolding of the phenomena which result from the insufficiency of cash holdings, we observe that in the first phase the fall in the general level of prices furnishes, by reducing the cash holdings required for the carrying on of transactions, the increments of cash holdings desired.

---

4. The rate does not depend upon the absolute level of prices, but only on variations in it. Mathematicians would say that it is a function of the derivative of the general level of prices with respect to time. (L'Ordre Social, Vol. I, p. 61.)

Now the rate of discount of the bank of issue is always very close to the market rate. When the rising market rate reaches the rate of discount, it stops increasing, since at this rate the bank accepts all paper offered and not demanded. From this moment on, all the excess of offers above the demand for short-term paper is diverted from the market to the bank of issue. The latter monetizes the paper which it has bought, and in this way supplies the increments of cash holdings desired. Commercial paper, however, is representative of wealth of the same value, wealth which is either stored up or, more generally, on its way through the process of production.

Everything goes on, therefore, as if the rights which contained this wealth, instead of being thrown upon the market, were disposed of outside the market in the assets of the bank of issue, the latter clothing them in the monetary garb which makes them reappear in the form of additional cash holdings.

Thus, in a regime of inconvertible money, as in a metallic regime, the non-utilization of certain incomes does not give rise to a lack of markets. Wealth of the same value as that not demanded is spontaneously diverted from the market to the bank of issue. There it is utilized for the manufacture of the increments of cash holdings demanded by the owners of the additional incomes which were not consumed and not invested. Thus, as long as the increase in cash holdings continues, the increments of production will find a market. The abstinence of the owners of the additional incomes will not have brought about under-employment.

The preceding analysis shows that in a regime of inconvertible money the process is analogous in principle, if not in form, to that characteristic of a metallic regime; but as a result of the great flexibility of interest rates, the first process is evidently more sensitive than the second. It will therefore act more easily and more promptly. In this way, it will assure a smoother adaptation of the productive apparatus to the opportunities offered it by the market. In a mixed regime — where money is obtainable both by the coinage of gold and by the monetization of commercial paper, the two processes may act simultaneously. The conclusion, from the point of view which interests us here, is not modified.

There is a case, however, where money and credit instruments do not represent wealth of equal value: when they are issued against engagements which draw their value only from a governmental

act, obliging the bank of issue to buy them at a nominal rate entirely different from that at which they could be sold in the market — the situation characteristic of every regime with a deficit financed by recourse to the bank of issue. In such a case, however, the rights which contain the false credits are added, when their owners wish to turn them into real wealth, to those from which the wealth offered on the market has been derived. The demand is increased in proportion. It is impossible, therefore, to imagine that the purchasing power impinging upon the market should not be sufficient to absorb the wealth offered there.

The preceding analysis shows that in all cases the demand for liquidity implies a demand for wealth of equal value. This wealth can, according to circumstances, be metal or credits, themselves representatives of goods stored or sold on credit. We are therefore not entitled to conclude that "liquidity preference" diminishes proportionately the purchasing power impinging on the market. This always remains determined, everything remaining the same, by the value of the production offered there. The demand for liquidity — like every demand, whatever its nature — simply sets forces in motion which tend to stimulate in the productive apparatus the adaptation capable of satisfying it.   .

To demand money is not, as Lord Keynes believes, to demand nothing. It is to demand wealth capable of being monetized within the framework of the existing monetary system. Hence, the preference for liquidity offers, like any other demand, an outlet for the labor forces offered on the market. Contrary to the Keynesian conclusion, it cannot be, at least so far as prices and the factors of production are not entirely immobilized, a cause of under-employment in the society which it affects.

III. The Origins of the Keynesian Error

The Keynesian theory of permanent under-employment equilibrium rests, then, essentially on an erroneous idea — the idea that all income not spent on consumers' goods or investment goods involves an inadequate absorption of the production of which it is the result. This idea is itself the consequence of two fundamental errors which characterize Lord Keynes' thought in the monetary sphere.

The first is based on the over-simplified idea that money and credit instruments are nothing but empty symbols with no value.

This, one might say, is the effect of a monetary nominalism with which the General Theory is thoroughly impregnated. The most characteristic passage from this point of view is the one dealing with financial provisions:

> But when the financial provision *exceeds* the actual expenditure on current upkeep, the practical results of this in its effect on employment are not always appreciated. For the amount of this excess neither directly gives rise to current investment nor is available to pay for consumption.
>
> Thus sinking funds, etc., are apt to withdraw spending power from the consumer long before the demand for expenditure on replacements (which such provisions are anticipating) comes into play; i.e. they diminish the current effective demand and only increase it in the year in which the replacement is actually made. (Pages 99, 100.)

Nothing shows more clearly that, for Lord Keynes, to accumulate reserves — that is to say, to accumulate money or short-term credit instruments — involves a proportionate diminution in the effective current demand, and therefore, the creation of under-employment.

The fallacy of this thesis appears immediately when the accumulation is in the form of metal. I have shown in the preceding section that the process characteristic of a circulation made up entirely of gold is general, and that the Keynesian thesis is just as untenable when the holdings consist of inconvertible money or short-term credit instruments.

At the beginning of Chapter 16, Sundry Observations on the Nature of Capital, our author presents the thesis even more clearly:

> An act of individual saving means — so to speak — a decision not to have dinner today. But it does *not* necessitate a decision to have dinner or to buy a pair of boots a week hence or a year hence or to consume any specified thing at any specified date. Thus it depresses the business of preparing today's dinner without stimulating the business of making ready for some future act of consumption. It is not a substitution of future consumption-demand for present consumption-demand — it is a net diminution of such demand.

Here there is no question that, for Lord Keynes, to save is to demand nothing. He does not realize that to accumulate money or credit instruments is to demand the values of which the money or credit instruments are a representation, and that to diminish one's cash holdings is to liberate the same values, causing them to be offered on the market.

The regulatory process thus neglected is, however, an essential one, indispensable to a comprehension of the monetary mechanism.

If it is not granted, it goes without saying that, as Keynes believes, preference for liquidity, that is to say, the accumulation of monetary reserves, tends to destroy the equilibrium of the market by inadequacy of demand, just as their utilization destroys it by excess. Every variation in reserves and holdings would, therefore, preclude the maintenance of economic equilibrium.

If, on the contrary, we grant it, the increase of reserves and holdings tends merely to divert to the fabrication of money the ' productive forces previously devoted to the production of the goods which are no longer demanded, while the utilization of these holdings tends to free the productive forces which were utilized for the production of the wealth represented by the money and orient them towards the production of the newly demanded goods

I believe, moreover, that the process of monetary regulation, if it is generally admitted so far as metallic money is concerned — though not always very conscientiously — is ignored by most monetary theorists, so far as inconvertible monetary systems are concerned. For my part, I have found it difficult to disentangle it and to show its generality.[5] I now believe it to be unquestionably established, and I believe, moreover, that it is the keystone of the whole theory of money.   `

In particular, I cannot see how one could explain the bond which must exist between the total amount of individual cash holdings and the quantity of money in circulation without making use of the theory of regulation. Every individual fixes freely, more or less consciously, the amount of his cash holdings. He generally ignores the existence of the procedures by which money can be created, and yet, in order that his desire for cash holdings may be satisfied, it is necessary that he be able by his decision to bring about variations in the quantity of money in circulation, in a regime of inconvertible money as well as a regime of metallic money. Only the theory of monetary regulation, based upon the mechanism which I have analyzed above, seems to me able to furnish the indispensable explanation and to show how each individual, in fixing the amount of his own cash holdings, helps to determine the total amount of money issued.

The problem of the bond between the amount of individual cash holdings and the total amount of money in circulation did not escape Lord Keynes, but since he ignores and denies the process of

5. *L'Ordre Social*, Vol. I, Chs. 17–20.

monetary regulation, he elaborates, to resolve it, an obscure explanation of the mechanism by which

the liberty, which every individual possesses, to change, whenever he chooses, the amount of money he holds [is harmonized] with the necessity for the total amount of money, which individual balances add up to, to be exactly equal to the amount of cash *which the banking system has created*. (Page 84. Italics mine.)

Thus, for Keynes, the quantity of money which the banking system has created is a datum. The total amount of individual cash holdings has to be adapted to it. I am convinced, on the contrary, that it is the total of cash holdings desired by individuals which, thanks to the mechanism of regulation, determines the quantity of money in circulation. But I have also shown that the mechanism of regulation, if we admit that it exists, excludes the possibility of equilibrium with under-employment and, consequently, destroys the foundation of the Keynesian theory.

It is not only the paragraphs which I have cited but the whole General Theory which leads to the conclusion that Lord Keynes' position is entirely dominated by the idea that the quantity of money in circulation is a datum arbitrarily fixed by the monetary authorities, upon which the market demand exercises no influence. His theory of interest (Chapter 13) in particular, rests upon this foundation:

It is the "price" which equilibrates the desire to hold wealth in the form of cash with the available quantity of cash —

And the quantity of money is not determined by the public. All that the propensity of the public towards hoarding can achieve is to determine the rate of interest at which the aggregate desire to hoard becomes equal to the available cash.

Furthermore, Lord Keynes believes that the monetary authorities can cause the quantity of money in circulation to vary. "If we have to govern the activity of the economic system by *varying* the quantity of money," he writes. Can a more outmoded and over-simplified conception be imagined? Among men who have reflected on monetary questions, are there any considerable number today who believe that a bank of issue fixes the quantity of money in circulation? All those who, from near or far, have participated with their eyes open in the management of a bank of issue are well aware that the open market can modify the cover of the outstanding circulation, can substitute to the great profit of

the bank, treasury bills for an advance to the state, and lower the rate of interest, but cannot directly modify the quantity of money in circulation.

As Directeur du Mouvement General des Fonds, I have known periods of equal deficit where the circulation increased and others where it decreased, without the monetary authorities having taken any action to bring about these changes and in spite of everything they could do to prevent them. As Deputy Governor of the Bank of France, I witnessed the vain attempts of the central bank to resist the increase of note issue.

Thus, the quantity of money in circulation, contrary to popular belief, is not fixed by the authorities of the market, and the fundamental error of Lord Keynes seems to me to result from the wholly superficial views which he holds concerning the monetary mechanism.

If we admit the existence of the mechanism of monetary regulation which I think I have demonstrated, under-employment cannot be a permanent state of equilibrium, since the mechanism of regulation tends to bring about those very transfers of the factors of production which are capable of making it disappear, even in the case where, as a result of liquidity preference, the demand of workers newly employed is exercised only in part upon consumers' goods and investment goods. Thus, either the quantity of money in circulation is a datum — and the theory of Keynes can be true — or the quantity of money is fixed by the size of the cash balances which the users of money desire to hold, and the Keynesian explanation of permanent under-employment equilibrium falls to pieces.

IV. The General Theory, Imperfect Philosophy of Unspecified Rigidity

It will be pointed out, to be sure, that the tendencies resulting from an increase of unutilized incomes, i.e. neither consumed nor invested, will not prevent unemployment unless they effectively divert productive forces from the production of wealth in the strict sense to wealth susceptible of being monetized, gold or commercial paper. Until the transfer has occurred, new production will throw upon the market wealth not demanded, and in this way will condemn to unemployment those who were disposed to devote themselves to such production. Under-employment will simply

be the expression of the refusal of the owners of incomes to accept what they do not want.

Thus, at the moment when the increase of employment takes place, if it is not directed into a channel which permits it to furnish the increments of money desired by the beneficiaries of the unspent increments of income, the situation may be that envisaged in, and explained by, the Keynesian theory. The only difference will be that while Keynes considers this situation a position of under-employment equilibrium, I regard it as a temporary state which the forces arising from the mechanism of regulation tend to modify.

If the action of these forces were rendered ineffective, how-ever, if they were incapable of diverting factors of production, Keynes' theory could then be considered a faithful explanation of reality. Thus, the theory of employment which Keynes calls "general" is valid only for very special cases, for economies which are entirely insensitive to movements of prices and of interest rates.

Moreover, in this case, if the theory is really to take account of reality, it would have to be the object of a profound generaliza-tion. If permanent unemployment can exist in an entirely rigid economy, it is not merely because the demand for investment goods may not be sufficient to offset the excess of a given increment of income above the increment of consumption which it is capable of causing, but because it might happen in various ways that the increase of production which an increase of employment might make possible in the channels where it is practically feasible would not consist of products which the beneficiaries of corresponding increases of incomes would wish to obtain.

An example will make my thought clearer. I assume a state of general under-employment; in other words, a state in which important segments of the labor force are either unemployed or employed less than they would like to be. Keynes says that in the absence of a systematic increase in investment this state might be a permanent state of equilibrium, because if employment increased, a part of the increments of income associated with the new produc-tion would not give rise to any demand, as a consequence of the psychological disposition of individuals to devote only a part — varying with their propensity to consume — of their increments of income to increased consumption.

I have shown that in such a case individuals who do not

consume demand money, and that the increments of cash holdings which their attitude leads them, consciously or not, to desire could be furnished them only by a suitable orientation of production — an orientation which the mechanism of monetary regulation tends to bring about. Hence, under the assumption which implicitly underlies Keynes' theory — a rigid economy and a propensity to consume not offset by an increase of investment — under-employment is permanent only because, while the owners of increments of income are not disposed to accept anything but increments of cash holdings, the increments of production which an increase of employment might afford them are only wealth in the strict sense — consumers' goods or investment goods.

Let the unemployed laborers apply themselves to the production of what is demanded, namely, gold in a country capable of producing it, exportable goods in a country possessing no gold deposits, or goods capable of being absorbed in a process giving rise to commercial paper, and employment can increase. Thus, in the Keynesian hypothesis, unemployment will result only from the incapacity of the productive apparatus to adapt itself to the market demand.

But this defect of adaptation — essentially temporary, since no one, it seems obvious, is disposed to hoard increments of income indefinitely — is only a very special and very exceptional form of the defects of adaptation possible. Under-employment is not caused only by an insufficient propensity to consume. It also results from every divergence between the increments of production which an increment of employment might supply and the increments of demand which the corresponding increment of incomes could give rise to.

Let us suppose, for example, that the situation in which Keynes sees the essential cause of under-employment does not exist, every owner of increments of income being disposed to demand consumers' goods for the totality of his new resources. Now in such a situation, where the propensity to consume would be 100 per cent, any increase in employment would be impossible if the workers capable of being newly employed were adapted only to the manufacture of investment goods or consumers' goods other than those which the owners of increments of income desired. The state of under-employment would exist despite a total propensity to consume.

Conversely, in a society where the unemployed workers were not ready to produce anything except consumers' goods — the case, in particular, of unemployed workers specialized in agricultural production — every demand for investment goods, however important, would have no effect upon employment. The Keynesian remedy would be wholly ineffective.

Thus, Lord Keynes has taken account, among all possible causes of under-employment attributable to economic rigidity, only one very special case, that of unemployment due to incapacity of the economic organism to furnish the increments of cash holdings or of short-term credit instruments which are temporarily demanded of it. He has given to this cause of under-employment an importance which it does not in general deserve, since unemployment can result from any defect of adaptation between production and the demand capable of absorbing it, and can last until this adaptation has been effected. He has, furthermore, failed to note that the complete economic rigidity required, if his theory is to be partially true, is not a general characteristic of economic societies but, on the contrary, a very exceptional state, which only special measures of immobilization or control could engender.

The omission in the general theory of the essential effects of economic rigidity is evidently an extremely serious matter, since it conceals the true character of the Keynesian explanation and brushes aside some of the remedies for under-employment which it should have suggested.

The considerations developed in the present section lead to a general view of the mechanism of unemployment. Contrary to Keynes' view, it does not result from an insufficiency of income. Income is never insufficient to absorb existing production; for, apart from special circumstances which I cannot consider in detail here, it is engendered by this production and its amount at every period is identically equal to the value of the said production.[6]

On the other hand, if the products offered are not those desired by the market, their value may be reduced to zero at the same time as the income of the producers to whose activity they are due. Thus the total income is not rendered incapable of absorbing the production, for the value of the latter is reduced in the same degree as the total of the former. But if the production of unwanted goods comes to an end, the state of unemployment to

6. L'Ordre Social — Ch. 10.

which it gave rise is not a state of equilibrium, for it engenders forces which tend to modify it with a view to restoring to the factors of production their normal productivity. It is only when these forces are systematically paralyzed that under-employment can become a permanent characteristic of the society in question.

V. The Political Consequences of the General Theory

The preceding analysis will enable us to form an opinion concerning the efficacy and probable consequences of the remedies for unemployment which the Keynesian theory suggests.

These remedies all rest upon the central idea that under-employment is due to an inadequate propensity to consume. To increase employment, therefore, it is sufficient either to increase the propensity to consume or to offset the inadequacy of the demand for consumption goods by a systematic increase of investment. In order to increase the propensity to consume, Lord Keynes recommends a redistribution of income designed to dis-discourage the deplorable instinct to save:

> Thus our argument leads towards the conclusion that in contemporary conditions the growth of wealth, so far from being dependent on the abstinence of the rich, as is commonly supposed, is more likely to be impeded by it. (Page 373.)

He also contemplates appropriate fiscal and interest-rate policies. If the level of investment is fixed, total income depends entirely on the propensity to consume, therefore on measures tending to develop it. "So long as the marginal propensity to consume out of wages is greater than that out of profits," says one of his disciples, "any rise in wage rates at the expense of profits will raise the aggregate marginal propensity to consume . . . thus raising the level of income that can be supported by a given level of investment and federal expenditure." (Econometrica, July, 1946, p. 227.)

The preceding analysis shows that these remedies cannot have any permanent effect on the level of employment, which is indifferent to the utilization made of the incomes to which it gives rise. It also shows that the corresponding interventions will reduce the temporary unemployment arising from economic rigidity only in the exact degree to which the increase in the propensity to consume arouses demand for the goods which the under-employed labor forces are capable of producing. If the latter are

unable or unwilling to offer anything but investment goods, the increase in the propensity to consume will leave them unemployed. In any event, their adaptation to the new markets afforded them by a given increase in the propensity to consume would be neither less difficult nor less painful than that which would have made possible the absorption of unemployment by adaptation to the utilization which the owners of incomes intended to make of them, whether it responds to a desire to save or a desire to hoard.

But the fundamental and quasi-universal remedy of the Keynesian theory is the investment expenditure undertaken by the state with a view to warding off the alleged inadequacy of private demand. For each level of investment there is supposed to be a corresponding level of income, and hence of employment. If employment declines, it is because the volume of investment required to sustain the existing employment has not been achieved. To do away with unemployment, it is necessary and sufficient that the state assume the investment expenditures which private initiative is unwilling to undertake.

The whole preceding analysis shows that this conclusion is false. The level of investment expenditures, whether public or private, does not define the level of employment, since with every level of employment there is associated an income capable of absorbing the corresponding production, under the one condition that the latter be adapted, in its nature, to the effective demand of the owners of incomes. Even if we admit "as a permanent characteristic of human nature" the existence of a consumption function analogous to that assumed in the Keynesian analysis, it does not lead to the conclusion that investment expenditures are necessary in order to insure full employment; for every demand which is not exercised upon the market for consumers' goods will reappear in the form of demand for investment goods or for hoarding.

It should be noted further that a demand for additional cash holdings will always be of limited amount, and that when it is satisfied, the corresponding demand will reappear upon the market for investment goods or consumers' goods.

It is true, however, that investment expenditures can bring relief to a temporary unemployment crisis, though only to a limited extent. They can furnish a market for unemployed labor forces available for the production of the investment goods for which they produce the demand. Every increment of investment

expenditure can increase employment in the investment industries and in these alone. Moreover, we should not consider the investment industries as a whole. It is only the factors of production speciaiized in the industries which benefit from the additional demand which will be afforded an additional market by the investment expenditures, relieving them from the unemployment which would have led them to make the adjustment required by the conditions of market demand.

However, though investment expenditures can in this way reduce temporary unemployment in the industries affected by them, they entail secondary effects which must be taken into account if we wish to arrive at a decision on balance concerning the consequences which the full-employment policy will bring in its train when it becomes the object of generalized application. These secondary effects·will vary according to whether the investment expenditures are achieved within the framework of a treasury in equilibrium or with a deficit; in other words, according to whether they are financed by taxes and loans or by the issue of treasury bills rendered eligible for discount because the market has not of its own accord assured the absorption of them.

In the first case, there is a levy on the society of the resources devoted to the financing of the investment program. If the purchasing power thus taken away from individuals is that which they intended to spend on wealth not offered in the market (for example, in the Keynesian hypothesis, additional cash holdings), and if the increment of demand arising from the investment program impinges on wealth which the unemployed factors of production are capable of producing, the investment program will increase employment, but it need not turn out this way. It is probable that in large measure the demand for articles not produced — for example, increments of cash holdings — will persist and that the levies accomplished will, to the extent of an important fraction of their total, reduce the demands which were impinging upon the other segments of the market.

Consequently, the program will have augmented the amplitude of the adjustments required for spontaneous reabsorption of the unemployed and delayed the moment when the latter will be able to come to pass. Under the (improbable) assumption that the public investment program has absorbed all the unemployed productive resources — that is, to the degree in which it has achieved

its purpose — it would have brought about the disappearance of every force capable of assuring an ultimate spontaneous recovery of the market.

But, furthermore, if the investment expenditures imply the utilization of raw materials or goods demanded in the market, they will, by increasing the demand and hence the price of these goods, help to reduce the outlets spontaneously afforded them by the market. So far as they have served to absorb this wealth, the investment expenditures will not have helped to increase the employment in the market. Finally, so far as the investment program diverts means of production from the areas where they are more desired to less useful employments, it will reduce the standard of living of the society.

However, it is unlikely that a large investment program following a period when economic depression has seriously reduced government revenues should ever be financed within the framework of a balanced budget. In the majority of cases, if not in all, resources will be obtained by the issue of treasury bills eligible for discount.

In the situation foreseen by the Keynesian hypothesis — a depression caused by the refusal of certain workers to utilize the increment of income afforded them by an increment of employment — inflation can supply them with the increments of cash holdings which they wish to obtain. In this way, so far as the offer of employment is accepted by the under-employed workers, whether it corresponds to their previous specialization or they accept the modifications in activity which it implies, an investment program financed by inflation can bring about an increase of employment.

However, individuals, other things remaining the same so far as prices go, cannot be supposed to increase their cash holdings indefinitely. The moment will necessarily arrive when the newly issued monetary tokens will not be wanted. Then they will produce, along with a rise in the general level of prices, all the economic and social disorders associated with inflation. If we wish to avoid the latter without abandoning the investment program which has given rise to them, there will be no other solution but to limit demand by a system of general rationing.

Thus, the inauguration of a vast program of public works, if it is carried out over a prolonged period, will revive in the world

an economic regime invented by Hitler, from which victory was supposed to free us. We shall see the restraints progressively tightening and expanding, and the steady unfolding of the familiar process of inflation will again bring about the suppression of all human liberties. In this way it will be demonstrated once more that the governments of human societies have a choice between only two solutions: to allow the apparatus of production to adapt itself to the structure which, by the movements of prices, the will of the consumers tends to impose upon it, or to adapt the desires of consumers by authoritative regulation to the structure of the productive apparatus which we do not propose to change.

The preceding analysis illuminates the phenomena which we have observed during the past decade and explains why the development of war industries caused unemployment to disappear, while investment plans applied in peace time seem incapable of accomplishing it. The war-time programs created a practically unlimited demand. They reabsorbed unemployment because the workers available were transferred, voluntarily or under compulsion, into the employments which this demand brought into being. As for financing, it was assured, so far as it was not covered by taxation or by loans, by recourse to the bank of issue. The inflation thus engendered was in large measure neutralized by rationing, that is, by the suppression of the freedom of the demanders in the utilization of their purchasing power.

The new activities obviously restricted the previous production by the utilizations of material and of energy which they implied, but no one thought of complaining about it, because at the same time taxation, borrowing, and rationing restricted the power of buying.

Can the same result — unsatisfactory as it is, since it implies and requires the suppresssion of all economic liberty — be hoped for in time of peace? I do not think so. In the first place, it is improbable that the administrative authorities will be able to impose in time of peace the transfers of labor power which such a program implies. These transfers will probably be neither less extensive nor less painful than those which would have assured the spontaneous reabsorption of the under-employed; and since the latter are considered unacceptable, it is improbable that the former would be any more acceptable, even if the public authorities had a mind to impose them. Moreover, the inauguration of a

large investment program will appreciably diminish, by the utilization of raw materials and of energy which it requires, the production of articles really demanded. Public opinion will be reluctant to give up what it wants for the production of what it does not want.

The privations which the investment program will cause will be much more appreciable than in time of war, for it will not be possible to raise the tax revenues to the level which they had attained during hostilities, or to obtain voluntary loans of such large amounts, or to impose, by means of rationing, a sufficient neutralizing of purchasing power. For all these reasons, an unsatisfied demand will persist in the market and this will give rise more or less rapidly, according to its relative magnitude, to all the troubles of inflation.

In spite of these prospects, it is probable that the next period of depression will see a general application in the world of the policy suggested by Lord Keynes. I am confident that this policy will not reduce unemployment, except to a very limited extent, but that it will have profound consequences upon the evolution of the countries in which it is applied. Through the economic disorders to which it will give rise, it will re-establish in the world a regime of general planning analogous to the regime of war time and based upon the suppression of all individual liberty. Thus, the next economic crisis seems likely to be the occasion for profound political changes, welcome to some people, dreaded by others. In any event, being based on a false theory, the remedies which will be adopted will give rise to repercussions very different from those they were designed to produce. Their ineffectiveness will be, for a great part of public opinion, one more reason for urging the suppression of a regime which, by denying itself, will have destroyed itself.

Whom Jupiter wishes to destroy, he first makes mad.

JACQUES RUEFF

INSTITUT D'ÉTUDES POLITIQUES, PARIS

THE BRITISH BALANCE OF PAYMENTS
AND PROBLEMS OF DOMESTIC POLICY[1]

SUMMARY

I. British postwar problems of adjustment, 368. — Possible avenues of escape from lower living standards, 372. — II. The government's double mandate, 374. — Comparison with the current American situation, 376. — Prospective capital shortage, 377. — III. Basic needs, 381. — Conflicting objectives, 383. — Resulting dangers: effect on exports, 386; on business incentives and voluntary savings, 388. — New outlooks, 390. — IV. The dilemma in terms of the balance of payments, 392. — Compromises inevitable, 393. — Conclusion, 395.

Most of the thinking in the United States in recent years about postwar economic problems has been concerned with maintaining a high level of employment. Even those economists who have concentrated on international economic reconstruction have been unable to exorcise the specter of another "American depression," much like Mr. Dick, in "David Copperfield," who could never keep King Charles' head out of his history of England. This preoccupation, which is essentially a preoccupation with the broader task of reconciling private enterprise with a greater measure of stability and security, is of course understandable after the experience of the 1930's.

The British have gone much farther than we in this sort of postwar planning. Both major parties are committed to comprehensive plans, many of which have already been enacted, to guarantee a basic minimum standard of living to all. This attempted guarantee will take two general forms: first, the provision of jobs, so that all may have an opportunity to earn an income;[2] second, a vast range of unearned benefits, consisting of supplements to income in the event of economic misfortune — unemployment, industrial accident, parenthood, sickness, old age, death — and expanded services provided directly by the government — slum clearance, housing, education, etc.[3]

1. The writer wishes to express his gratitude to Professors Myron W. Watkins and Howard S. Ellis, and Mr. Joel B. Dirlam for their suggestions and criticisms of this paper.
2. See the Coalition Government's White Paper on Employment Policy, Cmd. 6527, May, 1944.
3. White Papers on Social Insurance, Cmd. 6550-51, September, 1944; National Health Service, Cmd. 6502, February, 1944 (the basis for the Health

However, the war, aggravating an unfavorable secular development, has imposed upon Britain a second and fundamentally less pleasant task: that of adjusting to a new and poorer position in the world economy. The components of this problem are by now familiar:

(a) dependence upon imports for the maintenance of accustomed living standards;

(b) impaired ability to command those imports, because of
  (1) a long run deterioration of the competitive position of many of her export industries,
  (2) spreading economic nationalism, and
  (3) during the war, a further loss of markets, impairment of industrial capital, estimated officially at £885 million (through 1944),[4] resulting from obsolescence and heavy wear-and-tear of equipment with maintenance lagging behind, liquidation of £1,118 million (slightly over one-fourth) of her foreign investments[5] and loss of others, and the rise of a giant competitor in the American Merchant Marine; and

(c) a vast new short-term foreign debt, in the form of £ 2,879 million of new foreign-owned sterling balances accumulated during the war,[6] which will have to be paid off, in whole or in part. The

Bill of 1946); Educational Reconstruction, Cmd. 6458, July, 1943 (the basis for the Education Act of 1944); and others listed by Political and Economic Planning (PEP), "Reconstruction Plans," Planning, No. 227 (November 10, 1944), pp. 18-19.

4. "Statistical Material Presented During the Washington Negotiations," Cmd. 6707, December, 1945, p. 13.

5. Ibid., p. 9. The proceeds of war-time sales and repatriations may be compared with Sir Robert Kindersley's estimate of total prewar foreign investments of £3,725 million only after adjusting the latter upward to include investments in Eire, to take account of the appreciation in the sterling value of dollar securities at the outbreak of war, and a general increase in market values. See N. Kaldor and T. Barna, "The 1943 White Paper on National Income and Expenditure," ECONOMIC JOURNAL, vol LIII (1943), p. 261.

6. Total foreign liabilities of the United Kingdom amounted to £3,355 million on June 30, 1945, and are doubtless a few hundred million higher today. The official British estimate of total external disinvestment during the war (loss of foreign assets plus accumulation of foreign liabilities) is £4,198,000,000, including, in addition to the figures listed above, a decrease of 152 million in gold and United States dollar reserves, and an "unallocated" disinvestment of 49 million. "Statistical Material," cited above, pp. 10-12. These figures do not include Lend Lease obligations, which have since been funded at $650,000,000.

question is: how can the British either avoid or hold to a minimum the decline in living standards which seems in prospect as a result of their diminished ability to command the foreign goods they need and the necessity of paying off a new foreign debt?[7]

Since the balance of payments reflects these changes in Britain's international position, this question has usually been posed as a problem of balance-of-payments equilibration:[8] can the British people remedy the prospective shortage of means of payment for discharging their foreign debts and buying all the foreign goods and services which they would ordinarily demand when fully employed? If so, how are they going to do it? A fifty-five year loan from the United States, temporary moratoria on short-term foreign debts, and stringent controls over imports, it is generally conceded, are all helpful and necessary, but no more than temporary expedients with which no one will be long satisfied — neither foreign creditors and exporters nor the British people themselves.

Actually, of course, there is no question of the "ability" of the British people to overcome this "deficiency." It will be overcome whether they will it or not. Equilibrium in the balance of payments, in the usual sense of an equality in the demānd for and supply of the means of foreign payment without need for net gold flows or net short-term capital movements, does not present a problem, per se.[9] Since no country's supplies of monetary gold and

7. It is most unlikely that Britain will be willing or able to hold more than a small percentage of these foreign funds, now that she is committed to the earliest possible removal of exchange controls. Her net short-term foreign indebtedness as of August 31, 1938, amounted only to £760 million (Ibid., p. 10). This figure, which roughly approximated the dangerously high short-term debt of 1929 (Royal Institute of International Affairs, The Problem of International Investment, London, 1937, p. 339), resulted mainly from the heavy capital flight to Britain of the middle thirties, rather than from normal accumulation of deposits with the world's banking center. The mechanism of the World Bank and Monetary Fund will further reduce the need of foreigners to keep sterling balances.

8. The author has discussed this problem elsewhere, Great Britain in the World Economy (New York, 1946), Chap. XV; see also Randall Hinshaw, "American Prosperity and the British Balance of Payments Problem," Review of Economic Statistics, Vol. XXVII (1945), pp. 1–9; Randall Hinshaw and Lloyd A. Metzler, "World Prosperity and the British Balance of Payments," Review of Economic Statistics, Vol. XXVII (1945), pp. 156–70; Donald F. Heatherington, "British Postwar Balance of International Payments," Foreign Commerce Weekly, August 25, 1945.

9. Much of the following discussion of balance-of-payments equilibration

foreign exchange, or line of short-term credit abroad, are inexhaustible, it follows inevitably that in the long run no country can buy more foreign goods, services and securities (and make unilateral transfers) than it sells (and receives in other ways.) This truism may be restated in such a way as to apply to any individual or grouping of individuals.

This long-run equality is maintained, fundamentally, through alterations in the ability of the peoples of the various regions involved to command foreign goods and securities — relative curtailment in countries whose nationals are trying to buy too much, relative expansion in countries whose nationals are buying too little. Such alterations arise, automatically, out of the disequilibrium itself, and are of a magnitude and duration necessary to sustain a long-run balance.[1] These corrective changes in real buying power may take any number of forms, in any combination. Balance of payments disequilibrium automatically gives rise to changes in the domestic money supply and alterations (cumulative or otherwise) in relative levels of money income and expenditure, as Ohlin, Iversen, and others have stressed Shifts in income levels may take the form either of changes in rates of earnings, or in levels of employment (either or both of which may develop spontaneously, as in the case of the unemployment and wage and profit deflation in Britain's export industries in the 1920's, or be induced by a deliberate bank policy to maintain the external value of the currency). Shifts in terms of trade, reflecting either changed domestic prices or alteration in exchange rates, usually accompany and reinforce these corrective changes.

All of these developments have the effect of reducing the command over foreign goods in the weak countries and increasing it in the strong countries.[2] The various forms of government

is a summary of various sections of the author's book, cited above. See especially Chaps. I–III, and pp. 275–76.

1. This is not to suggest, of course, that the autonomous mechanism is at all times painless in its operation, or even that it provides the least painful method of reëquilibration available. The only real balance-of-payments problem is not: will equilibrium be achieved? but In what way and at what cost will it be achieved?

2. It is now commonplace that the terms of trade may·actually shift either way, although the probability, it is generally conceded, is that they will shift in favor of the country with the strong balance of payments. An opposite movement may indicate or create new disequilibria. In any event, the possibility (which is indeed the rule) of incessant disequilibration and reëquilibration of balances of payments does not invalidate our exposition.

intervention with the purpose of preventing or correcting disequilibria (it is not contended that all government intervention has this intention or this result!) operate in the same manner: whether by deflating, imposing import or exchange controls, subsidizing exports, or permitting the foreign exchange rate to fluctuate, governments alter the ability of people in one country to buy the goods, services or securities of another.

The anticipated "gap" between British expenditures and receipts, it may safely be predicted, therefore, will never materialize. It will either take care of itself or be taken care of by the government, in ways which will enforce a reduction in the real income of the British people — whether the reduction takes the form of import controls, depressed wage rates, a depreciated pound, or all three. But it is not the prospect of balance-of-payments disequilibrium in itself which is alarming; what is alarming is this reduction in real income which threatens. The answer can be found only in an attack on the causes of the prospective "gap": the world's unwillingness or inability to buy British goods, Britain's continued need for foreign goods, and the loss of wealth which Britain has suffered as a result of the war.

There are three possible avenues of escape from the threat of lower living standards. One is an increase in national efficiency, i.e. national productive capacity, which will make possible increased exports and, to a much smaller extent, a supplanting of foreign goods in fair competition in the home market.[3] The second is shrewd bargaining: in an exchange economy, real income is a function not only of productivity but also of the degree of control over the market, in short, of the terms of trade. The British government might attempt, as in the 1930's, to exact discriminatory advantages for British goods in foreign markets under the threat of excluding foreign exporters from the British market. By continued bulk purchasing, also, it might supply domestic needs at favorable, monopsonistic prices. The third, and most hopeful, alternative is a rapid rise in world income and trade, leading spontaneously to a more rapid increase in foreign demand for British goods than of British demand for foreign goods. Such an eventuality could conceivably solve the balance-of-payments

---

3. To the extent that home products supplant imports only "unfairly," e.g. as a result of a tariff, domestic living standards will still be bearing the cost of maintaining equilibrium.

problem painlessly, without need for any absolute decline in British incomes. For equilibrium requires only that Britain's command over foreign goods be depressed *relative* to the ability of the rest of the world to buy Britain's goods. It seems reasonable to expect an absolute secular increase in the latter from now on; without it the British prospect would be gloomy indeed.

The importance of the third possibility would seem to rule out reliance upon the second. "Shrewd" bargaining, in the form of government trading and discriminatory bilateral agreements, would invite retaliation and lead to trade wars, whose losses might soon exceed the benefits.[4] By intensifying the economic nationalism which had such a devastating effect on British trade in the interwar years, it would jeopardize the far more striking curative possibilities of rising world incomes. In such circumstances, even granting rising incomes abroad, no such stimulus to British exports could be anticipated as in a freely trading world. The record and outcome of beggar-my-neighbor policies is of sufficiently recent memory to require no summary. For Britain the record is clear: between 1929 and 1937, according to League of Nations indexes, world industrial production showed an increase of 20 per cent, and the output of primary goods 10 per cent, but the volume of world trade and of British exports declined three per cent and 17 per cent, respectively.[5]

Crucial as this third possibility is, we may exclude it henceforth from our consideration in this discussion of British domestic

4 This is not to say that the British government may be expected quickly and voluntarily to forego the immediate advantages of such bilateral negotiations, as long as it is responsible for the national food supply and for the conservation of exchange, and must in any case plan for the gradual settlement of war debts In the circumstances, such agreements as those recently concluded with Canada and Argentina, and other similar arrangements in process of negotiation, are understandable However, this *is* to say that to attempt to resurrect world trade on such a bilateral basis is to sacrifice long-run national necessities for temporary advantages.

A recent attempt to assess Britain's bargaining power, arising out of the dependence of foreign producers upon her market, and consequent ability to exact special concessions, concludes that it has been exaggerated, and offers little ground for optimism concerning the outcome of a trade war with the United States. G D A MacDougall, "Britain's Bargaining Power," Economic Journal, Vol. LVI (1946), pp 27–37

5. League indexes from World Production and Trade, British indexes from the Board of Trade Journal. The volume of exports alone is of course not an adequate gauge of national welfare, but in Britain's present circumstances it is an extremely significant one.

policy. For one thing, although Britain will, of course, have a considerable influence on the kind of world economic order which emerges and the chances for an expanding world trade, the fact remains that British policy alone will not determine the outcome. Moreover, there seems little prospect that world trade will revive so rapidly as to eradicate completely Britain's special balance-of-payments problem. Over the long pull, British exports will respond to world recovery in proportion to the ability of British industry to meet foreign demands more efficiently than its competitors. Between 1913 and 1929 the volume of world trade increased 27 per cent, according to League of Nations indexes, yet the volume of British exports declined 13.4 per cent.[6] Therefore — if our catalogue of Britain's "avenues of escape" is exhaustive — the problem will remain, in greater or lesser degree, one for domestic policy to solve. The British people must direct themselves to the first means of averting the decline in living standards (compared with 1938) which threatens or of permitting the recovery which is possible — the increase of national productivity. It is with the implications of this task that this paper is primarily concerned.

II

The British government, then, has a double mandate — perhaps it might better be described as a divided mandate — guaranteeing economic stability, security, and the basic minimum, on the one hand, and encouraging technological renovation, industrial adaptation to changed world market conditions, and the restoration of national wealth, on the other. The one springs from an aspiration typical of advanced industrial countries. As people have become better educated, better supplied, and better organized by virtue of the economic progress of the last century, they have sought government guarantee of a decent living standard for all and protection against the contingencies of a speculative economy. Unable by individual insight and planning to escape the vicissitudes inherent in a business enterprise system under present-day conditions, and unwilling longer to tolerate them, they demand of the state security of livelihood, as the price of their continued acquiescence in the capitalistic regime. The other mission is in essence but a special case of the perennial and uni-

6. Ibid., and League of Nations, Memoranda on Production and Trade.

versal economic problem, the need for economizing, arising from the inadequacy of means at hand for satisfying insatiable human wants. The one is an offshoot from a supposed "economics of abundance," based upon the fact of an already great productive capacity: it carries with it a rather pleasant prescription calling for realization of the potential plenty, in part by a redistribution of benefits, to maintain effective demand. The other springs from the implacable economics of scarcity: it suggests a none too palatable prescription for economic reconstruction and development, calling for expansion of productive capacity, increase of output, and repayment of debt; in a word, for frugality and hard work.

It is not suggested that to undertake these two tasks at the same time is to work at cross-purposes. Their goals are not necessarily and at all points incompatible. On the one hand, depressions mean a terrific loss in the efficiency of the productive system. Their elimination will mean greater output, and increased national productive capacity, providing the means of short-run stabilization do not interfere with the long-run adjustment processes which spell progress. On the other hand, the pressing tasks of industrial reconstruction will for a time avert depression. When there is ample work to be done, job security presents no problem.[7] Moreover, technological advances can be counted on to bring nearer and nearer the better life for all implied in the basic minimum. Even for the present, neither experience nor analysis affords much ground for doubting that capitalism can continue to provide adequate reward for productive effort, while at the same time being forced by government action to offer more security and benefits to the underprivileged than it would otherwise provide.[8]

7. See p. 389, below. The leading British "planners" are aware of this fact, but argue that the problem of unemployment for lack of effective demand will undoubtedly recur, perhaps ten or twenty years from now, when the present urgent consumption and investment needs are partially satisfied. The planning mechanism which they propose is intended to be a flexible one; for the postwar reconstruction period they conceive that it will probably have to be used selectively to restrain effective demand (excessive consumption, less important investment), in order to ensure the satisfaction of the more urgent demands — for reconstruction and the elimination of "Want, Disease, Squalor, and Ignorance." See William H. Beveridge, Full Employment in a Free Society (London, 1944), pp. 31, 146–59, 348, 368, 390–93; PEP, "Manpower Stocktaking," Planning, No. 253 (August 2, 1946), p. 13.

8. Unquestionably there are limits beyond which, at any given time, social security legislation can not go without weakening the mainsprings of

Nevertheless, it is true of these two prime subjects of British solicitude that the fears to which they give rise, the psychologies which they engender, and the policies which they suggest are indeed in large measure contradictory and incompatible. A Labor government, pledged to strive for both full realization of the economics of abundance and all-out attack upon the economics of scarcity, may find its efforts (and accompanying promises and appeals) in one direction dissipating those in the other. More specifically, it may jeopardize the chances of resolving the second, more basic, difficulty by its concern with the first.

Current American experience offers an imperfect analogy to this British dilemma. The point of departure for much economic speculation in this country during the war was fear of the deflationary consequences of a cessation of military spending. Yet in the event we have faced inflation, at least during the first postwar year and a half. Policies suitable for the prevention or correction of the one are, of course, not applicable to the prevention or correction of the other. The crucial problem of forefending or combating either is one of timing.[9] But the slowness with which a three-compartment democratic government acts, and the incomplete conversion of Congress to the viewpoint that it must assume responsibility for an active policy of economic stabilization (which calls for something more than a pious expression of intention, as in the recent Full Employment Act) necessitate both prediction and propaganda, if appropriate policies are to be adopted in time. The limitations and dangers of prognostication and agitation as the basis for economic policy are manifest. Many economists, proclaiming the danger of imminent deflation in 1945–46, partly for propaganda reasons, encouraged wage and fiscal policies which have proved inappropriate.

Moreover, in the circumstances, the public — which in any case finds it difficult to understand that what may have been correct yesterday, and what may be desirable tomorrow, may be disastrous today — is easily confused. Broader social security coverage, higher unemployment compensation may be desirable

enterprise (on the part of either the beneficiaries or those who bear the costs), but these limits seem as difficult to define as capitalism's limit of tolerance to an internal public debt. The whole question of economic incentives stands in need of further exploration and more incisive analysis.

9. See Committee for Economic Development (CED), *Jobs and Markets* (New York, 1946), Chap. VIII, *passim*.

in the long run, yet undesirable today.[1] Higher hourly wage rates in 1947 may indeed help prevent a depression by providing a mass market for a reconverted and expanded industry, yet during 1946 they contributed to inflation. Tax reductions may soon be imperative, but would be inexpedient today — and so on.

The British dilemma is more serious, because it is more fundamental: theirs is not a matter of the proper timing of policies to offset short-run business fluctuations, but a basic "clash of progress and security," to use the dramatic characterization of the dilemma by Professor A. G. B. Fisher, in a book of that title.[2] That issue bids fair to perplex them for many years.

Modern technology, embodied in modern capital equipment, has been the mainspring of economic progress. A continued capital shortage is the prime characteristic of the "economics of scarcity." Britain — indeed the world as a whole, outside the United States — appears to have entered upon a new period of relative capital shortage,[3] which should persist for the next ten or even twenty years. This situation is in part the direct result of the war: recon-

1. Tens, possibly hundreds, of thousands of veterans drew unemployment compensation during the summer of 1946 despite the fact that they were students on vacation, and had not the slightest intention of working. These payments contributed to inflation by discouraging the taking of jobs, and by further inflating the money income stream.

2. New York, 1935.

3. As long as wants are unsatisfied, there may be said to be an inadequacy of capital. In this sense, the capital shortage of Britain or India was scarcely less acute (save for the intervening war-time depreciation and destruction) in the 1930's than it is today. What then differentiates the prospective "capital hunger" of the next decade or two from the "capital satiety" of the 1930's, to borrow Colin Clark's terminology? (The Economics of 1960, Toronto, 1942, Chap. VII–VIII.)

In the earlier period the prospective returns from capital formation, as viewed by the business man, who bore the prime responsibility for investment decisions, were such that the demand for savings was far less, even at the then low interest rates, than the supply that would have been forthcoming at full employment. For the next few years, investment, conducted in greater measure than heretofore by governments, will probably match, if it does not exceed, the supply of voluntary (ex ante) savings. This will create upward pressures on national incomes and interest rates, and lay the basis for a rise in living standards by increasing productive resources.

In a perfect market there could be no persistent capital shortage greater in one area, for example "outside the United States," than elsewhere. But political and psychological impediments to foreign lending by the United States, based on a distrust of the willingness or ability of foreigners to repay (a distrust which may prove well-founded because of a perverse American commercial policy) may perpetuate this paradox of capital shortage elsewhere, capital satiety here.

struction and reconversion of productive facilities, rebuilding of devastated cities, harbors and bridges, continued military expenditures and insistent pent-up consumption demands all vie in making demands on the limited supply of (real) savings.

Moreover, we may be witnessing the upswing of a cycle of investment activity in the development of economically backward areas of the world. One sign pointing in this direction is the social awakening in those areas long dormant; witness the political upheavals in India, Indo-China, Indonesia, China, the Middle East, Eastern Europe, Argentina, Bolivia, Venezuela. But cyclical forces would probably have started an upward thrust of investment by now even without the galvanizing action of war. The drought of international investment in primary producing areas during the prewar decade would in any event probably have resulted in a relative recovery of the prices of food and raw materials during the 'forties. A new wave of investment, such as occurred between 1900 and 1929, was evidently in the cards.[4]

Be that as it may, certainly the peoples of the Far and Near East, South America, and Eastern Europe will not be content to return to primitive, one-crop cultivation. They are intent on the introduction of modern methods of cultivation and the development of local industry. Before they are through, they will have revolutionized their economies. In the process they will require capital. Many of these countries already have the means at hand in war-swollen dollar and sterling balances; use of the latter will be a direct drain on British savings.[5] Finally, an improvement in the competitive position of British industry will require complete renovation of antiquated plant and machinery in the staple, export trades. Output per man-hour in Britain was far below that in the United States, even before the war, and is undoubtedly farther behind now.[6]

4. Such a relative recovery of primary prices around the turn of the century, following upon a decade of greatly reduced international investment, undoubtedly stimulated the heavy investment that followed. See the author's book, cited above, pp. 144–49.

5. India, Burma, and the countries of the Middle East alone accounted for no less than £1,732 million of the £3,355 million of short-term external liabilities of the United Kingdom on June 30, 1945. "Statistical Material," loc. cit., p. 11.

6. See Philip S. Brown, "Prospective National Income and Capital Formation in the United Kingdom," American Economic Review, Vol. XXXVI (1946), pp. 555–556, and the author's book, cited above, Chaps. V–VII.

Since, even if they wanted to, the British people cannot hope to attract from foreign sources more than a very small portion of the capital which they require, they must provide it out of domestic savings. But unless national income increases more rapidly than anticipated, it seems unlikely that they will voluntarily forego consumption to the extent necessary to accomplish the tasks which need doing.[7] Statistical "computations" of a nation's income and outlay in the future are at best of dubious worth, but what computations have been made bear out this judgment.

For example, projections made by Nicholas Kaldor in 1944[8] yielded estimated national savings of £575 million in 1948 (all estimates are at 1938 prices) and consumption outlays (public and private) of £5025 million out of a national income of £5600 million. This volume of savings would permit "only . . . a modest increase in the rate of capital expenditure in industry and in housing," according to Kaldor. Highly tentative formulations of two alternative investment plans more nearly approximating Britain's needs (in 1948) yielded figures of £750 and £1000 million — 31 per cent and 74 per cent, respectively, in excess of estimated voluntary savings.[9] The larger figure provides a hypothetical £140 million for net investment in plant and machinery, compared with £45 million which Kaldor suggests might be available if national savings amount to only the estimated £575 million.[1]

These figures seem reasonable in the light of a few more recent specific plans and estimates of capital requirements for the rehabilitation of the three industries most in need of drastic overhauling. Direct comparison may not be made, however, because these figures are in current pounds sterling, which are of considerably smaller value than the 1938 pounds in which Kaldor's estimates are expressed, and because these reconstruction expendi-

7. See Mary E. Murphy, "Trends and Conflicts in the British Economy," *American Economic Review*, Vol. XXXVI, Papers and Proceedings (May, 1946), pp. 628–41.

8. Beveridge, op. cit., Appendix C.

9. Ibid. The figure of £750 million includes provision for 500,000 houses annually ("regarded as the necessary rate of building, if an adequate number of houses are to be provided, and all slums abolished, in a period of 15 years") and a 92 per cent greater investment in fixed capital (including public works) than actually occurred in 1938 (in real terms). The second estimate includes a still larger investment in plant and machinery and public works, and a contribution of £150 million to the reconstruction of Europe

1. Both plans make additional provision for "other fixed capital" and public utilities.

tures are planned for a period of years, whereas Kaldor's figures are for annual outlays.[2] A recent plan submitted to the government by the British Iron and Steel Federation suggests an expenditure (over a period of years) of £168 million for that industry alone — a sum which the London Economist characterized as "certainly not too great." Capital expenditures by this industry amounted to one or two million annually in the late 1920's and about £6 million annually a decade later.[3] For mechanizing the coal industry, the recent nationalization legislation provides for expenditures of £150 million in the next five years, and promises more if necessary.[4] The capital costs of reequipping the cotton textile industry have been put as high as £250 million.[5]

Only Kaldor's larger figure for over-all net investment requirements makes any provision (£150 million for 1948) for a renewal of international lending, to assist in the reconstruction of Europe and to rebuild depleted foreign assets. Yet so great is Britain's dependence on a revival of world trade that such a contribution is most advisable, despite the country's own urgent needs. Moreover, though these estimates assume a sufficient production of export goods to close the gap in the current balance of payments, neither makes any provision for debt repayment. If the blocked sterling balances are funded and paid off at the same rate as the American loan, these payments for services already rendered will amount to an additional £100 million annually for the next 50 years. On the other hand, the American loan will make available £187,500,000 annually for five years, but thereafter will impose an added burden of £30,000,000 a year.[6]

2. British wholesale prices were in the spring of 1946 some 72 per cent above their 1938 level; the cost of living had risen roughly 31 per cent in the same period. The estimates of required capital expenditure in these specific industries, which may refer to gross investment, may on this account also not be closely comparable with Kaldor's hypothetical figures for net investment outlay.

3. "Plan for Steel," The Economist, Vol. CL (May 11, 1946), pp. 759-60.

4. "Coal Bill Inquest," The Economist, Vol. CL (May 18, 1946), pp. 805-6.

5. "A Policy for Cotton," The Economist, Vol. CL (January 5, 1946), p. 23.

6. The American loan is to be repaid, beginning in 1951, in 50 equal (except for the last) annual installments, at the rate of $31,823,000 for every billion dollars of credit used. Since the credit will presumably make possible an expansion of Britain's productive capacity sufficient to create the means of paying the (interest) cost of the loan, the interest will represent no net new

In a recent article, Philip S. Brown has concluded that Kaldor's projections were unduly optimistic. By projecting pre-war trends, Kaldor assumed a 13 per cent increase in man-hour productivity between 1938 and 1948. Brown, calling attention to the war-time depletion of capital equipment and the decline in that part of the population which is of working age, a decline in immediate prospect, concludes that man-hour productivity will be no higher than a decade earlier, although an increase of perhaps 14 per cent may be anticipated by 1952. His estimate of net capital formation, even in that later year, is only £520 million, at 1938 prices, compared with Kaldor's 1948 estimate of £575 million.[7]

III

In a country as exposed as Britain to the vagaries of international competition — particularly one which has just lost perhaps 25 per cent of its national wealth[8] — neither economic security nor a better livelihood yield to direct search. For the nation as a whole, the former can be attained only as the by-product of technological progressiveness, of constant adaptation and change; the latter, similarly and equally paradoxically, can come only as the by-product of abstention. Britain must have a high rate of saving, to restore the lost wealth and increase national productive capacity. Taxes will have to remain high, to pay off foreign debts and the cost of expanded social security, free public education, and slum clearance. Moreover, high-bracket incomes alone cannot provide these revenues. For one thing, these sources are inadequate. For another, progressive taxation reduces the supply of savings. Hence Britain cannot afford quickly to reduce

burden, but the repayment of principal will require a substitution of British savings for American credit.

Repayment of the Lend-Lease settlement of $650,000,000 and of £3 billion of sterling balances (£5 million and £95 million annually, beginning in 1951, if settlement of the latter is on the same terms as the former), payments for services already rendered, will represent in toto a net new burden on the national product.

7. Op. cit., American Economic Review, Vol. XXXVI (1946), pp. 557-70.

8. This official estimate is based on a comparison of estimated external disinvestment, £4,198 million, domestic disinvestment (accumulated arrears of depreciation and obsolescence of business property only), £885 million, physical destruction on land and sea, £2,200 million, with estimated national wealth of £30,000 million. "Statistical Material," loc. cit., pp. 12-14.

the heavy war-time direct taxes on the middle and lower ranges of income and indirect taxes on consumer goods and services; these levies yield large revenues to the Exchequer while at the same time tending to check the propensity to consume and to import.[9] Continuation of import controls, rationing of basic foods, and allocation of raw materials into export products and essential reconstruction are other unavoidable necessities in such a program of "austerity," at least until the balance of payments has righted itself sufficiently to make possible relaxation of such controls. Other measures of national abstinence must include a halt to money wage increases which necessitate price increases, to enable British exporters to compete, and to restrain the propensity to consume, and sacrifice of restrictive provisions in collective bargaining contracts, a perfect example of a direct quest for job-security which jeopardizes the technical progressiveness from which alone truly social security may be found. For the time being, the British people may even have to lower the sights for the housing program, because housing uses savings (and hence labor, materials, and foreign exchange) which may better be used elsewhere. It may be difficult to accept the idea that any use of scarce resources could better serve the public welfare than the provision of adequate housing; in fact, in their present circumstances, if the British people have to decide whether to divert some lumber from housing to the construction of temporary movie "sets" for films which may please foreign audiences, they may be wise to do so, for upon the recovery of exports rests the hope for more and cheaper food.[1]

9. To the extent that the revenues are paid out as social security and similar benefits, such taxes hold down the increase of the national propensity to consume.

1. Again, one may well question the advisability of continued import of such foreign luxuries as American motion pictures, if it becomes necessary to curtail residential construction in order to conserve foreign exchange, and to encourage the exports which make possible such relatively frivolous imports. Continued limitation of such imports will be unavoidable for the time being. However, the British government may have to think twice about relying solely on this means of enforcing national "abstinence," if these trade controls jeopardize the success of a co-operative international program for the lowering of trade barriers. See pp. 386–388, below.

At the same time, the British have a right to demand as quid pro quo for freeing from governmentally-imposed restraints foreign films seeking access to the British market that the United States free its domestic market from such private trade barriers as the control of distributive channels by American producers, which is equally effective in limiting the ability of foreign films to compete freely for the consumer's favor and money.

Yet in weighing all these possible courses a labor government is likely to be torn between its two mandates. Social security, higher old age pensions, broader education, including the raising of the school-leaving age from 14 to 16, and government-subsidized housing unfortunately run up against the economics of scarcity. All entail a higher national propensity to consume, a using up rather than a conservation of savings.[2] Already rising wage rates, pressure by unions for guaranteed annual pay and other contractual welfare provisions, predicated in most cases upon anticipated rather than already accomplished increases in productivity, have led to higher unit costs. Thus the benefits (in terms of expanding exports without increasing imports pari passu) of the prospective integration, mechanization, and modernization of such industries as coal, steel, textiles, have been partially dissipated, in advance. The lack of strict national wage controls is politically understandable; so are the much higher wages decreed for coal in hope of holding and attracting reluctant workers to the mines; so also is "the most powerful cost-raiser of all . . . the factor that cannot be registered in statistics — the universal reluctance to do a hard day's work to which almost every employer of labour could at present testify," reported by the London Economist.[3] But understandable or not, the results are likely to be serious.

In the circumstances, the influence of direct provisions for security upon economic incentives can not be ignored. Consider, first, the effect on the incentives of those who bear the taxes. No one can prove the existence of, much less locate, the limit of tolerance of business and investors to graduated income taxation. Yet

2. This is true as long as the taxes levied to cover the costs of these programs diminish to the slightest extent the savings of the taxpayer, as of course they must. Social security payments involve a higher national propensity to consume, if the beneficiaries save less of the payments than the taxpayers would have saved; this, too, is of course probable. The Economist dramatically expounds this dilemma in "Assumption D," Vol. CL (February 2, 1946), pp. 161–62.

3. "Costs and Efforts," Vol. CL (January 19, 1946), p. 81; see also "The Risk of Inflation," Vol. CL (May 25, 1946), p. 827. The liberal weekly New Statesman and Nation reports the same attitude, with the return of the "atmosphere and the traditions of peace," and lays the blame on the failure of the labor government to give workers new incentives to hard work. "Incentives in Transition," Vol. XXXI (April 20, 1946), pp. 275–76. Meanwhile, government leaders themselves are apparently seriously concerned over "go slow" tactics, tolerated, if not encouraged, by unions. New York Times, November 6, 1946, p. 1.

there must be some limit, and it seems certain, for the time being at least, that any attempt to meet the expanded demands of the Exchequer by heavy taxes on higher incomes and corporations alone, while reducing taxes in the middle and lower income brackets to prewar levels, will no longer suffice, and may seriously impair entrepreneurial incentives and the supply of savings. The problem of the incentives of those who are protected also deserves probing. Barbara Wootton suggests that the only ultimate solution to modern society's dilemma of unwillingness, on the one hand, to let people lack minimum provision for health and decency — an unwillingness which will soon take concrete form in British law, while remaining largely verbal in the United States — and, on the other, fear of withdrawing the threat of starvation as an economic incentive is to provide such a minimum for all, while at the same time providing more than that for those who do work.[4] Unfortunately, today this remains, for England, a counsel of perfection. In the meantime the risk remains — a risk no doubt worth taking, but a risk none the less — that broad plans of social insurance, "more jobs than men" (or even as many jobs as men), and standard wage rates or levels of income not closely related to effort or productivity — all in the interest of security for the individual — will so weaken the will to work, and hence so raise unit costs as to threaten wholesale insecurity.[5]

The danger of excessive attention to security is not confined to wage earners. Britain has suffered in recent decades a progressive ossification of business enterprise, the evils of which are obvious in an economy which depends upon business initiative for continuous adaptation and steady growth. The story of this declining virility is a familiar one. Many signs in the last quarter of the nineteenth century indicated that German business men, in particular, were quicker to introduce improved methods of production and more aggressive in finding and holding customers than their British competitors.[6] The problem became far more acute in

4. Freedom Under Planning (Chapel Hill, 1945), p. 101. Beveridge offers a similar caveat and a similar solution. Social Insurance and Allied Services (New York, 1942), pp. 6–7.

5. See "The Carrot and the Stick," Economist, Vol CL (June 29, 1946), pp. 1033–35, and, on the other side, Beveridge, Full Employment in a Free Society, pp. 194–98.

6. Ross J. S. Hoffman, Great Britain and the German Trade Rivalry, 1875–1914 (Philadelphia, 1933), Chap. I and passim; Alfred Marshall, "Memo-

the period between the two World Wars when, under the stress of increasing adversity in international competition, excess capacity, and the manifest need for industry-wide reorganizations, encouraged by a sympathetic government and a new tariff protection, British business leaders systematically eradicated the spur of competition at home and abroad.[7] The results of this sheltering of inefficiency are apparent throughout British industry, not merely in the basic export trades — coal, iron and steel, textiles — but even in the newer industries, where the spirit of competition was weak from the outset, and has been further impaired by patent controls and widespread cartel agreements with foreign producers.[8]

Nationalization and the preparation of comprehensive "plans" for inefficient industries, the remedies proposed by the Labor Government, do not automatically ensure correction of these deficiencies. This truism has been repeated frequently in recent years, and deserves continued reiteration. Nevertheless, such measures attest to an intention to attack the problem of industrial inefficiency vigorously, a determination to make the necessary capital available, and an unwillingness to permit a respect for vested business interests to prevent the weeding-out of the laggards. These are certainly favorable signs, but the task has yet to be done.

André Siegfried chided Englishmen in 1931 for their inability to tighten their belts and get to work. An excessive appreciation of the pleasures of modern civilization, he claimed, unfitted them in the struggle for existence. Their refusal to retreat from accustomed standards of living, their resistance to wage reductions and hard work, their acceptance of a dole as a matter of right, had the result that

England is trying to compete in international markets and at the same time provide her people with a wage level and a standard of living which does

randum on the Fiscal Policy of International Trade," Official Papers (London, 1926), pp. 405–6.

7. Arthur F. Lucas, Industrial Reconstruction and the Control of Competition (London, 1937); Ben W. Lewis, Price and Production Control in British Industry, Public Policy Pamphlet No. 25 (Chicago, 1937).

8. In automobiles, rayon, electrical products, and various branches of the chemical industry, for example, evidence accumulated of a British lag behind competitors in the interwar years. See the author's book, cited above, Chap. VII. On British participation in cartels in the chemical, electrical and

not permit costs to be low enough either to export profitably, or to attract the capital necessary. . . .

At the moment when they should be redoubling their efforts, when everything has to be reconstructed on a new pattern, they prefer to rest on their laurels.[9]

During World War II the British, faced with the inexorable necessity of economic sacrifice, rose to the occasion. It is more difficult to realize that the economics of scarcity may impose a similar requirement in time of peace, even though the war was a "victorious" one. Realization of this need and appropriate action must be particularly difficult for a people who have experienced rising real incomes for a century. Because of their continually rising productivity, and because of the rapid increase in foreign demand for their goods, the British have found it possible in the past to obtain steadily rising wages for declining hours of work, and increasing social services. No inherent defect in the British character, such as Siegfried implied, prevents them from surmounting their current difficulties. Invidious speculative comparisons with the response of other nations to similar circumstances would be interesting but irrelevant. The fact remains that in Britain's straitened circumstances, continued pressure for higher wages, for governmental benefits supported either by progressive income taxation or by taxes on payrolls, in expectation of (and in advance of) increase in productivity and in world demand for British goods, may prove dangerous.

The dangers are two. First, high wage costs, whether imposed by union pressure or payroll taxes, will hamper British exporters in international competition, once the extreme scarcities of the immediate postwar years are eliminated. Since world demand for British goods will soon be very elastic again, the effect will be an aggravation of the balance-of-payments disequilibrium. It may be argued that in this event all that would be required would be a decline in sterling, which would eradicate both the competitive handicap and the advance in real incomes of workers entailed by the higher wages. However, under the Bretton Woods mechanism, quick approval for substantial alterations in exchange rates will not be easy to obtain; nor is any British government

other fields, see George W. Stocking and Myron W. Watkins, Cartels in Action (New York, 1946).

9. England's Crisis (New York, 1931), pp. 132, 177; see also pp. 27-28, 91-106, and passim.

going to be anxious to resort to systematic devaluation of sterling as an offset to rising costs. On the contrary, unless a sudden decline in American prices occurs, the British are likely to cling to the present artificially high rate of $4.03,[1] just as they did during the war. Competitive depreciations were to be expected in the prewar era of underemployment; overvaluation, accompanied by selective direct controls, has greater appeal in an era of scarcity. Considering the inelasticity of British demand for foreign goods, and the fact that expensive raw material imports mean high export costs, no British government is likely to choose poorer terms of trade to avert gaps in the balance of payments. Hence excessively high export prices are likely to mean continued import restrictions, rationing and allocation of scarce materials, all of them not merely vexing and embarrassing to a country which is seeking to co-operate in freeing world trade of restrictions, but also in the long run a threat to British living standards, which depend on an expanding world trade.

1. It is impossible to determine conclusively whether, let alone how much, sterling is overvalued or undervalued compared with the dollar. Purchasing power parity rates, based upon some period in the past when balances of payments (and hence, presumably, exchange rates) were apparently in equilibrium, have never been an adequate criterion, because relative price levels are not the sole determinants of the transactions which constitute the balance of payments. Canadian prices rose more rapidly than British between 1900 and 1914; this price change was neither a cause nor an indication of a progressive over-valuation of the Canadian dollar, but on the contrary was a result of the great increase of investment activity, accompanied by a rapid influx of capital from abroad — which bolstered rather than weakened the Canadian dollar.

So the present dollar-pound rate apparently reflects fairly accurately the relative rise of British and American wholesale prices in the last ten years. Professor Howard S. Ellis has informed the writer of a recent computation by the Board of Governors of the Federal Reserve System which yielded a parity rate as of May, 1946, of $3.8645, taking the period October 1936–July 1937 as base. The recent rise of American prices has doubtless corrected the slight divergence indicated between parity rate and the present market rate of $4.0336. Nevertheless, sterling may in the interim have become overvalued, because Britain has since become a debtor, and faces the need for making heavy unilateral payments to foreign holders of sterling balances, while having lost important sources of income from abroad. And a future decline in British compared with American prices, implying a progressive undervaluation of sterling, may be nothing more than an equilibrium adjustment to the requirement of a higher ratio of British exports to imports of goods and services than in 1936, if equilibrium is to be maintained.

Of course, price parities mean even less than ever today, in view of the widespread domestic and foreign trade and exchange controls. The use of purchasing power parities rests on the assumption of a free market: more

The second danger is that wage and tax costs will encroach excessively upon business incentives and the supply of voluntary savings, for both of which the need is exigent in an economy which continues to rely primarily on private enterprise. Employees in capitalist economies have in the past obtained wage increases and the benefits of social legislation, not primarily out of "surplus value," but from the fruits of technological progress, powered by the quest for profit and fueled by a flow of saving. The Labor Government unquestionably recognizes, as have reform governments in other countries, that such gains cannot with impunity and without limit be extracted from the rewards for enterprise and investment and out of incomes which would otherwise be saved, without jeopardizing the advancing national productivity which alone makes possible the continued increase of such benefits.

These considerations suggest the need for a revaluation of the rôle of savings in the British economy in the next decade or two. High wages, social services for the needy, financed by progressive taxation, have long been justified on grounds of equity. Out of the experience of the past twenty-five years has emerged a second justification: that such disbursements protect "mature" economies (or, as may prove to have been the case, economies in the "mature" phase of a long cycle of investment activity) against permanent or recurrent unemployment A tendency for savings at full employment to outrun profitable investment outlets is a cardinal article in this doctrine. Progressive taxes and high wages have been urged by postwar planners, therefore, as a means of sustaining effective demand. People are exhorted, not to parsimony, but to greater

precisely, it assumes a substantial area where marginal units move solely in response to price changes, and it assumes no change in the freedom of the market during the period under consideration.

Actually, as the writer has argued elsewhere (op. cit , pp 171–73), the entire concept of "overvaluation" or "undervaluation" of a currency is misleading, implying as it does that the balance-of-payments disequilibrium, or the need for trade and exchange controls, symptomatic of an incorrect exchange rate, was *caused* by the divergence of that rate from an equilibrium level, or that the correction of that divergence is the proper way of eliminating the disequilibrium

The statement that sterling is overvalued today therefore means only that, if Britain were to remove her controls over prices, imports, and consumption, and to make blocked sterling balances freely convertible into foreign currencies, the pound would decline compared with the dollar. If this is not true, then there is no British balance-of-payments problem.

consumption, in order to take full advantage of the potential abundance about them.[2]

With the re-enthronement, in Britain, of the economics of scarcity, it becomes clear that, for some time at least, excess savings will be conspicuous by their absence. The problem will be to maintain an adequate rate of saving. This is true, not only because of the pressing demands already noted, but also because the higher the propensity to save — the more the British people abstain voluntarily from consumption, whether of foreign goods or of domestically produced goods and materials which might otherwise be exported — the less will be the need for trade controls to equate receipts and payments, and the more therefore can Britain contribute to loosing the bonds on world trade in general. For both these reasons, the less the British consume today, the higher will be their real income tomorrow.

The widespread fear that savings will be inadequate, that consumption expenditures (individual and social) will be too high, suggests that the British have little cause to fear large-scale unemployment in the next decade or so. This seems definitely to be true. If it is, the demands for security can in large measure be satisfied without enforcing any further redistribution of income, and without interfering with the rehabilitation of British industry and the replacement of lost wealth.

The General Staffs of the Allied countries have frequently been accused of having spent the interwar period, as usual, preparing for the war which had just passed, rather than for that which was to come. Economic planners may be forgiven an occasional glance at the past, for whereas the technology of warfare moves constantly "forward," the problems calling for economic planning are in large measure recurrent. Even if no one business fluctuation is exactly like any other, there is the indisputable fact of a cyclical movement. Nevertheless, the British postwar planners, concentrating their attention on the problem of full employment, may well be subject to the same criticism as the generals. It is well, of course, to be prepared for a deflation fifteen years hence or, perhaps, for one sooner, spreading eastward from the United States, which is in the happy position of really having to plan to

2. See, e.g., Beveridge, Full Employment in a Free Society, pp. 95–96, 157, 165–66, 186.

make full use of its potential abundance.[3] But if such planning is based upon a distrust of saving, and intentionally or unintentionally encourages even steeper progressive taxation, higher wage demands (to "maintain purchasing power"), job-security provisions in labor contracts, inflexible prices, and an avoidance of competition, it may prove positively detrimental in the altered circumstances of today.

There are other signs of a recognition of these altered circumstances besides the emergence of a new (rather a recurrence of the traditional) evaluation of savings.[4] A declining birth rate, like the quest for security, is a product of higher living standards, and has a similar motivation. Political and military considerations are largely responsible for the alarm with which the declining Western birth rate is currently regarded. Also, in the last ten years, some economists have argued that capitalism, under which we have traditionally directed a substantial portion of available productive resources to providing for ever-larger future generations, will now be unable to provide adequate employment opportunities because of the altered market prospect.

However, more recently the prospects of an economy of scarcity have reawakened fears akin to those of the Mercantilists regarding the economic consequences of a stationary or 'declining population. For example, the influential Political and Economic Planning organization (PEP) has recently characterized the danger as follows:

> By 1949 or 1950 we shall have nearly 18 millions as our civilian working population — rather less than those available in 1939. Thereafter the age distribution of this population will become more and more adverse to production as the population as a whole grows older, and later on the total numbers are bound to shrink. With those human resources Britain has got to restore and maintain home consumption, to build up a larger volume of exports . . . and to add again to this volume in order to repay her overseas indebtedness.[5]

3. "Our capacity to produce is enormous — perhaps 40 or 50 per cent more than our actual 1939 output. Failure to consume and invest at rates far higher than prewar will mean mass unemployment." CED, op. cit., p. 113.

4. Sir William Beveridge, himself, while reciting several times the dampening effect of savings on effective demand, stated clearly that consumption would probably have to be restrained in the reconstruction period. See p. 375, note 7, above.

5. "A Programme and a Purpose," Planning, No. 246 (March 15, 1946), p. 6.

Recently Prime Minister Attlee has spoken in the same vein:

> We are faced with a shortage of manpower. We must see to it that it is used to the best advantage, and that means a changed attitude of mind. For years before the war we were accustomed to having surplus labour on the market, to having a large amount of unemployment, and the existence of that surplus labour bred in all classes an attitude of mind that must be changed. . . . Therefore, all classes must get away from the old ideas, and realise that we are now going to live in a world of full employment.[6]

The assumptions underlying these statements are not the assumptions behind the economic homilies of the 1930's. They evince no fear of an excess labor force. They carry no suggestion that Britain may have become incapable of supporting its expanded population at accustomed standards of living.[7] Nor do they demonstrate any fear that a stable or declining population threaten unemployment, for lack of investment opportunities. The basic assumptions are both anti-Malthusian and "anti-Hansenian": the implicit predicate is that a progressive technology can put off the point of diminishing returns as the British population grows.[8] The attitude is simply that with so much work to do a chronic shortage of manpower threatens to delay completion of the task of recon-

6. Quoted by PEP, "Manpower Stocktaking," Planning, No. 253 (August 2, 1946), p. 1. The alarm caused by the prospect of a declining population is not confined to reformist groups and parties. It was the Coalition Government, led by Conservatives, which set up the Royal Commission on Population, which has for two years been conducting a family survey to ascertain the causes of the low birth rate. New York Times, May 8, 1946, p. 17. See also Brown, op. cit., pp. 557–59.

7. There is no way of determining the number of inhabitants a given region can support, at given standards of living, at some time in the future, as the author has argued elsewhere. "Palestine: A Problem in Economic Evaluation," American Economic Review, Vol. XXXIV (September, 1944), p. 559. However, the decline in the competitive position of British export industry and the march of economic nationalism in the interwar years gave rise to justifiable fears that the British Isles could no longer support a population which had grown and prospered as a highly specialized part of a network of expanding world trade.

In part, the altered attitude toward the population prospect may be attributed to general post-war optimism. In part, however, it is undoubtedly attributable to the return of the economics of scarcity — to the fact that once again there appears to be more than enough work to be done.

8. See, e.g., PEP "Population — A Challenge and a Choice," Planning, No. 251 (June 14, 1946), p. 7 and passim. At the same time this pamphlet, without recognizing any implicit contradiction, does make a concession to the Keynes-Hansen viewpoint, stating as a second cause for alarm the threat of unemployment for lack of investment outlets. pp. 7–8. However, the authors advance this as definitely a secondary consideration.

structing British industry and raising living standards. Though this is a recurrence of the economics of scarcity, it has a counter-Malthusian twist: it is inadequate population, not an excessive one, which carries the threat of poverty.[9]

IV

Since the altered circumstances which are responsible for these changes in attitude toward savings and population growth result primarily from an altered British position in the world economy, it may be helpful once more to state in terms of the balance of payments the dilemma facing the British government. A prospective deficiency of means of payment on international account resulting from a deteriorated competitive position, loss of foreign investments, and accumulation of a large foreign debt, will tend to bring about a decline in British compared with foreign purchasing power. In the absence of government efforts, this decline might take one of two, or a combination of two, alternative forms: a depreciation of sterling exchange, and a relative deflation of money prices and incomes at home. If the latter, it might be imposed by heavy taxes raised to meet foreign obligations, or by an unwillingness of foreigners to buy expensive British goods. Since all of these more or less automatic methods have undesirable aspects, the government has already taken a hand, and undoubtedly will continue to assume responsibility. It may help to ensure, or expedite a return to, equilibrium by maintaining import controls, by rationing and other limitations on consumption, by deflationary monetary and credit policies, or by controlling money

9. Here again we find some carry-over of the attitudes of the 1930's. The speech of the Prime Minister and the PEP pamphlet which quotes it with approval imply at times that the sole occasion for "Manpower Stocktaking" is the government's determination to ensure full employment, which will create a "seller's market" for labor. Planning, No. 253, p. 12. It is true, of course, that the fuller the employment governments strive to underwrite, the nearer they set their sights to a point where there will be more jobs than men, the greater will be the need for centralized controls, priorities, and manpower budgets. See Melvin G. de Chazeau, "Employment Policy and Organization of Industry After the War," American Economic Review, Vol. XXXV September, 1945), pp. 630–35. However, as PEP at other times recognizes, at least for the next ten years, and possibly even longer, it will not be the Government's full employment policy which will create manpower shortages, but the fact that there is so much work to do. See "Manpower Stocktaking," loc. cit., p. 13, and "Population — A Challenge and a Choice," loc. cit., passim.

wage rates. Wage controls would contribute to reëquilibration by making possible lower export prices, and by restraining the propensity to consume (freeing scarce goods and materials for export, and reducing imports). At the same time, they would diminish the need for other restrictions which impede the reduction of barriers to world trade.

The question, then, is not: will equilibrium be achieved? Rather, it is: how may it be achieved least painfully and in the best interests of the British people in the long run? The only solution, ultimately, is to restore the lost wealth — to pay off war-accumulated foreign debts, make up arrears of plant depreciation and obsolescence, reconstruct bomb-devastated areas, and even, if possible, to refill the nation's portfolio of foreign investments.

Restoration of the national wealth is the key to sustained improvement in the competitive position and productive power of British industry. To the extent that the British people maintain an equilibrium in external trade between national intake and outlay by restraining wage increases and checking consumption expenditures (by voluntary savings, and mitigation of excessively high progressive taxes), rather than by retention of import controls or by depreciation of sterling in the exchange market, they will at the same·time be working toward a more permanent solution. They will be helping British goods to win their way back in foreign markets, as well as supplying the savings which are necessary both to make the monetary transfers and to modernize and reëquip British industry. Industrial reconstruction and growth alone can provide the higher living standards, the high real wages, and the security which they seek.

In practice, by combined force of political pressure and economic necessity, the government will inevitably compromise. And insofar as the twin goals of security and progress are not utterly antagonistic, probably no other course than compromise and pragmatic experimentalism would be desirable. The British people will demand and, judging from their current mood, obtain legislative provision for greater social security, subsidized housing and broader education, in short, a more equitable distribution of income. By inexorable economic necessity, they will not gain them without a substantial levelling of living standards among the working classes themselves. They will obtain housing for which their need is so desperate, and gradual relief from the continued

shortages of the more common necessities — but neither as quickly nor in the abundance they desire.

Various government controls — over imports, over the allocation of scarce materials and capital, hence over consumption — imposed both to conserve foreign exchange and to encourage production for export and industrial renovation, will enforce this continued abstinence. The government may also have to continue efforts to stimulate savings, for the economic emergency is not yet over.[1] There will have to be similar compromises in the field of international policy — between bilateral trade and monetary negotiations, such as have been made during the last year, under pressure of balance-of-payments stringency and the need to settle accounts suspended during war, on the one hand, and, on the other, multilateral efforts to reconstruct the freest possible world trade.

Trade unions will obtain various benefits for their members, but they will also have to submit to some limitation on their contract provisions which restrict output or the introduction of technological improvements. They will probably encounter governmental pressure (as well as the pressure of brutal economic reality) to prevent continuous wage increases, except as these prove to be justified by rising productivity. Incidentally, the far better strike record in Britain than in the United States since V-J Day, and the attack on restrictive practices by the Prime Minister and union leaders at the recent Trades Union Conference[2] suggest that British labor may prove more coöperative than American labor in all these respects, whether because of their greater faith in a Labor government, or because of the prospective guarantee of security against unemployment (which eliminates much of the pressure for job-security provisions), or because of the longer history of collective

1. The Coalition Government's White Paper on Employment Policy states that "the habit of saving must still be encouraged" during the "transition from war to peace," loc. cit., Chap. II, par. 16. The fact which needs stressing is that in terms of the continued need for a high level of saving, the "transition" is likely to be of longer duration than merely the period of physical reconversion from war to peace production. See, e.g., the statement on Employment Policy and Organization of Industry after the War, drawn up at Oxford University, Nuffield College (London, 1943): "There will be *short-run* problems of readjustment connected with demobilization and the transfer of industry from a war to a peace basis; but beyond these looms the more permanent problem of 'full employment' after the transition has been made." Par. 1 (italics supplied).

2. New York Times, October 25, 1946, p. 8.

bargaining in Britain. On the other hand, the election of a Labor government does not automatically eradicate labor problems, as the Minister of Fuels and Power can attest today[3] And such a government, devoted to a program of insuring full employment, may find it particularly difficult to exercise even moral pressure to hold down wage increases, although it is certainly attempting to[4] and will have to continue to do so for many years.

It is an unfortunate consequence of the peculiar problems which Britain faces that after six years of privation her people must still postpone in some measure the good life which they seek. However, the dilemma is one which the as yet unresolved economic problem of scarcity — a problem aggravated by the colossal

3. The serious lag of coal output during the past year has been the result, primarily, of the inability of the government to recruit labor for the mines, or to keep those working in the mines from absenting themselves from work with discouraging frequency. See, e.g., "Desperate Remedies," The New Statesman and Nation, Vol. XXXI (June 29, 1946), p. 461. Indeed, one effect of the close coöperation between union leaders and the government has been an increase in the number of unofficial, wild-cat strikes, since the Labor Party's victory. New York Times, August 31, 1946, p. 6.

4. While the government has followed the general policy of accepting the decisions reached by collective bargaining (see, e.g., "Britain's Wage Problem," Labor and Industry in Britain, published by British Information Services, April 1946, pp. 66–68), it has stressed over and over the primary importance of keeping costs low while expanding output. As the extremity of Britain's problem has become increasingly manifest, the government's injunctions against wage increases have become increasingly direct. In its "Statement on the Economic Considerations Affecting Relations Between Employers and Workers," Cmd. 7018, January 1947, the government nowhere categorically states its opposition to wage increases unjustified by demonstrable increases in productivity, although such an opposition is implied throughout the White Paper. On the other hand, in the more recent, and more gloomy, "Economic Survey for 1947," Cmd. 7046, February 1947, published in the midst of the coal crisis, the government warns that "any further general increases in wages and profits must be accompanied by a corresponding increase in production" (p. 9). The inclusion of profits, which are not a serious problem in the circumstances, might appear to the skeptic further evidence of an unwillingness to face up squarely to the problem of wage increases. A statement concerning the length of the work week, later in the same report (p. 31), is even less equivocal: "The nation cannot afford shorter hours of work unless these can be shown to increase output per man-year. Greater leisure is a very desirable thing, but it is not at the moment a prime essential like imported food."

The reader will note that the door is still left open for wage increases and reduction of hours of work. The coal miners have recently obtained a five-day week, for example, on the promise that absenteeism will decline and output will not suffer. However, under pressure of emergency conditions, the government has clearly been forced to adopt a firm stand.

stupidity of war — poses for most of the peoples of the world. The British people may find some comfort in the certainty that this problem is constantly in process of resolution. What is called for is not a permanent reduction in consumption, but a more gradual, better sustained, and, in the not-too-long run, a greater increase than might otherwise occur. The mere attainment of full employment in itself lessens the need for abstinence. The more rapid reconstruction, renovation of industry, and technological progress which such restraint will make possible will after a gestation period not overlong create steadily rising per capita incomes. With them will come a simultaneous rise in savings and living standards. Progressively, barring another relapse into the savagery of war, the more agreeable problems of making full use of potential abundance will reassert themselves. Intelligent planning is as necessary today to assist in speeding that progress as it will be ten or twenty years hence to assure the full enjoyment of its fruits.

ALFRED E. KAHN.

RIPON COLLEGE

SOME PROBLEMS
IN THE EXPLANATION OF INTEREST RATES[1]

SUMMARY

I. Scope of the paper, 397. — The credit market, 397. — Model A, 400. — Model B, 403. — Model C, 403. — Model D, 404. — Income lags, 405. — Ex-ante and ex-post concepts, 406. — The determinants of active savings, 409. — The determinants of depreciation allowances, 411. — The determinants of credit creation, 412. — Purchase and sale of securities, 412. — The demand side: new investment, 413; reinvestment, 414; cash holdings, 415. — Length of the elementary period, 417. — Keynes on the classical authors: savings, 418; net flow and gross flow, 420. — II. The markets for long and short credits, 423. — The demand for short credits, 424. — The supply of short credits, 425. — The demand for long credits, 426. — The supply of long credits, 426. — The cyclical behavior of interest rates, 427. — Interest rates in the principal countries: differences in level, 432; characteristics of movements, 432. — Econometric analysis, 433. — Interest rates and the price level, 434. — Physical determinants of the interest rate, 436. — Explaining the cyclical movement of interest rates, 437.

I

1. The present paper embodies a digest of some recent literature on the theory of interest rates, with the special purpose of using that theory for econometric analysis. The final purpose of that analysis may be formulated as the explanation of the differences between: (a) interest rates in different countries at one time; (b) various types of interest rates in the same country and at the same time; and (c) the development in time of these interest rates. All econometric analysis has to start from a theoretical scheme of the subject in question; and the object of such a theoretical scheme is to indicate the "determinants" or "data" of the rates of interest and the functional relationship between these rates and their determinants. These determinants or data must be entities outside the market complex on which the interest rate is formed.

2. The market complex in question is the credit market.[2]

1. I am much indebted to Prof. Dennis H. Robertson for valuable comments on the first draft of this article.

2. I prefer in this respect the standpoint taken by Ohlin, Alternative Theories of the Rate of Interest, Economic Journal XLVII (1937), p. 423, Haberler, Prosperity and Depression, 3d edition, 1941, and Lerner, Alternative Formulations of the Theory of Interest, Economic Journal, XLVIII (1938), p. 211, to that taken by Keynes, although I consider the difference unimportant.

For an explanation of interest rates we have to analyze demand for and supply of credit. The element of this market, the single credit contract, has two dimensions at least: a certain amount is lent for a certain time period and the total of all credits outstanding at a given point of time is a so-called "ensemble renouvelé," like the human population. In order exactly to describe this "population" we have to use rather complicated symbols.[3] We shall indicate by $_TK_t$ the total amount of new credits started during the elementary period t and expiring at T; $_TK_t^o$ the total amount of outstanding credits at the beginning of the elementary period t and expiring at T; K_t^n the total amount of new credits started during t, irrespective of their period of expiration; hence K_t^n is obtained by adding up all $_TK_t^n$'s for various values of T:

$$K_t^n = \Sigma^T {}_TK_t^n. \tag{2, 1}$$

Similarly, the total amount of credits outstanding at t

$$K_t^o = \Sigma^T {}_TK_t^o. \tag{2, 2}$$

Indicating by K_t^E the total amount of credits expiring during t — which, by the way, equals the sum $\Sigma_T K_t^n$ taken over all values of $T<t$ — we have the following calculation for the amount of credit outstanding at the end of period t:

$$K_{t+1}^o = K_t^o + K_t^n - K_t^E \tag{2, 3}$$

which is quite analogous to the calculation of a population at December 31 from the figure at January 1, plus the number of births, minus the number of deaths.

I could conceive of another description of the market complex that is relevant for the explanation of interest rates, viz. by taking the market complex for all elements of wealth, including e.g. the stock market which, strictly speaking, is not a credit market in the narrower sense and nevertheless exerts considerable influence on the interest rates; in fact, the yield on stocks may be said to represent one of the interest rates. I shall not, however, follow this line of attack.

3. The system of notation followed is based on the following rules. The *right-hand* index at the *bottom* indicates, for flows, the elementary time period and, for stocks, the time point to which it relates. The beginning of an elementary time period is indicated by the same number as that period and hence the end point by the next higher number. The left-hand indices at the bottom are used to indicate sorts of commodities (consumption goods: C; investment goods in general: I; investment goods for reinvestment: R; for new investment N, etc.), or sorts of credits (long: L, short: S) or the time of expiration of a credit: T. Left-hand top indices are used for the denotation of sorts of households spending (not those receiving) a given type of money amount (family households: 1; firms: 2; banks: 3; all households on the demand

The period of expiration of a credit need not be a fixed one; it may depend on several, yet unknown, factors. It may, for example, depend on the market conditions prevailing at a later time, as in the case of mortgage bonds, which may be bought back by the issuers at a time point at their own convenience. In such a case, the individual credit can only roughly be classified, and the symbol $_TK_t^n$ is indeterminate; K_t^n, however, remains determinate. The same situation exists in the case of a bond which changes hands. Let a four per cent bond be issued by firm A and be redeemable after forty years, the subscriber being Mr. B. Provisionally a credit for forty years is yielded by B to A. If, however, after five years, Mr. B sells his bond to Mr. C at a price of 80, the credit of B to A expires and is replaced by a new credit of C to A, with an interest rate equal to the yield at that moment which will be somewhere between four and five per cent.[4]

By K_t^N we shall understand the increase in credit outstanding, i.e.

$$K_t^N = K_{t+1}^o - K_t^o = K_t^n - K_t^E. \qquad (2, 4)$$

The use of this concept will appear in section 19, where we shall discuss the "stock-and-flow antithesis."

Demand for and supply of credits relate to the new credits started K_t^n, including, therefore, inter alia sales of old bonds on the stock exchange.

3. The essential feature of a credit to the debtor is disposal side of a market: D, on the supply side: S, etc.). The right-hand top indices are reserved for various indications, as e.g. sorts of income (interest: I, rent: R, etc.), outstanding vs. new credits (O, N, n, respectively; see below), etc. Summarizing:

Left-hand top: spending households		Right-hand top: mixed indications	
Family	1	Credits outstanding	O
Firms	2	New credits (gross)	n
Banks	3	Credits expiring	E
Demand side	D	New credits (net)	N
Supply side	S	Interest income	I
Left-hand bottom: goods		Right-hand bottom:	
Reinvestment	R	Time	t
New investment	N		
Long credits	L		
Short credits	S		
Time of exp.	T		

For simplicity's sake, if no confusion threatens, sometimes indices are omitted.

4. At the same time, B suffers a capital loss, which means a transformation of part of his credit to A into a pure gift.

of means of payment or money; hence the close relations between
the demand for and supply of credit and of cash. A credit may be
granted (i) by transfer of existing means of payment from one
subject to another, and (ii) by the creation of new means of pay-
ment. This creation is supposed to be the special function of the
banking system. In order to get, for a given model of the national
economy, a complete survey of all credit transactions, it is con-
venient to start with a survey of all money transactions and then
to distinguish between two sorts of transactions: credit and "other"
transactions. The "other" transactions may be characterized,
from the viewpoint of the recipient, as receipts from present or
past activity and from the viewpoint of the payer as payments
"out of his own means."

The models we are to consider will be very much simplified
models of a national economy. Throughout this paper we shall
restrict ourselves to a closed national economy; we shall not
consider the implications of international transactions, since this
would involve us in too many new problems. In all our models
we shall distinguish between only three sorts of households:
(1) family households, (2) firms (excluding banks, but including
public authorities), and (3) banks, to be indicated, where necessary,
by top indices 1, 2, and 3 at left. We assume vertical integration
throughout the system, i.e. we consider all business firms as a
unit. For the same reason we abstract from the original supplier
of and the last demander for credit we shall consider. Of course
there is room, in more detailed studies, for the introduction of the
intermediary markets as well. The only remaining function of
the banks in this picture is the (positive or negative) creation of
credit.

Looking first at that part of the credit market that consists
of the transfer of existing means of payment, we may therefore
restrict ourselves to the first two types of households. In each
category there will be households with a net supply of and house-
holds with a net demand for credits. It is impossible to say before-
hand whether a given household belongs to the one or the other
type; that may depend on the rate of interest itself. This does
not matter very much, since the algebraic sign of the expression
"demand for minus supply of credit" will automatically indicate
whether there is a net demand or a net supply.

4. In our first and simplest model, "model A," we assume

inter alia that (i) all business profits are paid out to stockholders (i.e. to family households); (ii) production is carried on without the help of durable means of production, implying that there are no depreciation charges; (iii) there are no financial transactions (i.e. no dealing in "old" securities, only the floating of "new" capital); (iv) there is no lag between income generated and income paid out.

The system of transactions of family households during the elementary period t then consists of:

Payments		Receipts	
Consumption Expenditure	C_t	Incomes	Y_t
"Active" Savings (or savings		Credits redeemed	K_t^B
handed over)	S_t		
Redemptions reinvested	K_t^B		

The introduction, on both sides, of the term K_t^B is a question of definition of S_t only; in practice it is impossible to distinguish between the reinvestment of redemptions and the investment of new savings; "net" savings handed over will be considered simply as the balance between gross savings handed over and redemption receipts.

Indicating the cash holdings at the beginning of the period by ^1M_t and those at the end by $^1M_{t+1}$, the following connection exists between these items:

$$Y_t - C_t - S_t = {}^1M_{t+1} - {}^1M_t. \qquad (4, 1)$$

This equation is equivalent to a definition of "active savings." By S_t we understand that part of income Y_t which is neither consumed (C_t) nor hoarded ($^1M_{t+1} - {}^1M_t$); hence:

$$S_t = Y_t - C_t - ({}^1M_{t+1} - {}^1M_t), \qquad (4, 2)$$

which is identical with (4, 1).

This definition requires two comments. First it must be noted that by hoarding we simply mean the increase in cash holding; this hoarding may be distinguished into intended and unintended hoarding; these notions will be discussed later (sections 9 and 11). Apart from these two types of "family hoarding," there may exist "business hoarding." Further, our definition of active savings is exclusive of hoarding; this implies a difference from both the Keynesian and the Robertsonian definitions of savings: Keynes includes all hoarding (in our sense) in his savings, whereas Robertson includes all intended (family) hoarding.

Savings may, for a number of households, be negative. This simply means that their contribution to S_t is a negative one; S_t taken for all households represents "net" active savings also in this sense. Of the transactions enumerated the term S_t and the double term K_t^E on both sides are the credit transactions, relevant for our analysis. As a rule, S_t will be positive and represent one of the supply items of the market equation.

5. The system of transactions of business firms consists of:

Payments		Receipts	
Income payments	Y_t	Sales of consumers' goods	C_t
Investments	I_t	Sales of investment goods	I_t
Redemption of credits	K_t^E	New credits taken	$^D K_t^n$

Indicating by ^2M_t and $^2M_{t+1}$ the cash holdings at the beginning and the end of period t, the relation holds:

$$Y_t + I_t + K_t^E + {}^2M_{t+1} = C_t + I_t + K_t^n + {}^2M_t. \qquad (5, 1)$$

This may also be written as:

$$K_t^n = K_t^E + I_t + (Y_t - C_t - I_t) + {}^2M_{t+1} - {}^2M_t. \qquad (5, 2)$$

If there is no lag between income generated $C_t + I_t$ and income paid out Y_t, and hence

$$Y_t = C_t + I_t, \qquad \qquad (5, 3)$$

(5, 2) becomes:

$$K_t^n = K_t^E + I_t + {}^2M_{t+1} - {}^2M_t \qquad (5, 4)$$

telling us that new credits are, then, needed for three different ends:

(a) to replace credits redeemed;

(b) to finance new investments; and

(c) to finance a desired increase in cash holdings.

As the non-credit items in the system of transactions we consider Y_t, C_t and I_t which, in our case, just cancel out. This is not so in the case of a lag between income produced and income paid out.

For the group of business firms as a whole, K_t^n as a rule will be positive and therefore represent a net demand for new credits.

Considering now the credit-market as a whole, we shall find that the net supply of credits by the family households S_t need not always be equal to the net demand for credits by business units $^D K_t^n$. The gap, if any, may be filled by the banking system in the form of positive or negative creation of credit M_t^N. Hence the equation expressing the balance between supply of and demand

for new credits will run:

$$K_{t}^{E} + S_{t} + M_{t}^{N} = {}^{D}K_{t}^{n}. \qquad (5, 5)$$

which may be specified with the aid of equation (5, 4):

$$K_{t}^{E} + S_{t} + M_{t}^{N} = K_{t}^{E} + I_{t} + {}^{2}M_{t+1} - {}^{2}M_{t} \qquad (A) \qquad (5, 6)$$

where the terms K_{t}^{E} may be left out altogether.

6. We shall now consider a number of complications of our model that will bring it nearer to reality. In "model B" let part of the national income take the form of undistributed profits ${}^{2}Y_{t}$; family incomes will then be:

$${}^{1}Y_{t} = C_{t} + I_{t} - {}^{2}Y_{t} \qquad (6, 1)$$

and family savings:

$${}^{1}S_{t} = {}^{1}Y_{t} - C_{t} - ({}^{1}M_{t+1} - {}^{1}M_{t}) \qquad (6, 2)$$

${}^{2}Y_{t}$ may be said to represent business savings and equation (5, 6) becomes:

$${}^{1}S_{t} + M_{t}^{N} = (I_{t} - {}^{2}Y_{t}) + ({}^{2}M_{t+1} - {}^{2}M_{t}) \qquad (6, 3)$$

The term $I_{t} - {}^{2}Y_{t}$ in the right-hand member now stands for that part of investment which is not financed out of business savings. Of course we might also bring over ${}^{2}Y_{t}$ to the left-hand side of the equation:

$${}^{1}S_{t} + {}^{2}Y_{t} + M_{t}^{N} = I_{t} + ({}^{2}M_{t+1} - {}^{2}M_{t}) \qquad (B) \qquad (6, 4)$$

where ${}^{1}S_{t} + {}^{2}Y_{t}$ again stands for total active (family plus business) savings.

Similar considerations may be applied for the case (model B') where part of family savings is invested directly, say in new house building. Indicating this type of investment by ${}^{1}I_{t}$ as distinct from business investment ${}^{2}I_{t}$, equation (5, 6) turns into:

$$S_{t} - {}^{1}I_{t} + M_{t}^{N} = {}^{2}I_{t} + {}^{2}M_{t+1} - {}^{2}M_{t} \qquad (B') \qquad (6, 5)$$

where $S_{t} - {}^{1}I_{t}$ now represents active savings as far as not invested directly. Also here the equation might be written in the form

$$S_{t} + M_{t}^{N} = {}^{2}I_{t} + {}^{1}I_{t} + {}^{2}M_{t+1} - {}^{2}M_{t} \qquad (6, 6)$$

where I_{t} appears now decomposed into two components.

7. A more important step consists in the introduction ("model C") of depreciation charges, as a consequence of the use of durable means of production.

It is natural, in this case, also to decompose I_{t} into two components: reinvestments ${}_{R}I_{t}$ and new investments, ${}_{N}I_{t}$. Income

Y_t now no longer equals $C_t + I_t$, but:

$$Y_t = C_t + I_t - N_t = C_t + _RI_t + _NI_t - N_t \qquad (7,1)$$

where N_t represents depreciation charges. Equation (5, 2) now runs:

$$K_t^n = K_t^E + _NI_t + _RI_t - N_t + {}^2M_{t+1} - {}^2M_t \qquad (7,2)$$

and the market equation (5, 6) for new credits becomes:

$$S_t + N_t + M_t^N = _NI_t + _RI_t + ({}^2M_{t+1} - {}^2M_t) \qquad (C) \quad (7,3)$$

Here there appear three sources of supply: net savings, depreciation charges and credit creation, and three origins of demand; new investments, re-investments and increased cash holdings.

8. One further complication worth considering is the introduction into the picture ("model D") of financial transactions, i.e. purchases and sales of "old" securities. Since it is difficult to distinguish between purchases of old and of freshly floated securities, it seems logical to speak of purchases of securities ${}^DE_t$ only, and to indicate by ${}^1S_t$, all other forms of active savings. In the same way, let us indicate by ${}^IE_t$ new issues of securities, and by I_t' all other credits taken for investment purposes. If, finally, sales of old securities are written as ${}^sE_t$, the family budget runs:

Payments		Receipts	.
Consumption expenditure	C_t	Income	Y_t
Purchases of securities	${}^DE_t$	Sales of securities	${}^sE_t$
Other active savings	${}^1S_t'$		
Increase in cash holdings	${}^1M_{t+1} - {}^1M_t$		

whereas the business budget is:

Income payments	Y_t	Sales of consumers' goods	C_t
Investments	I_t	Sales of investment goods	I_t
		Capital issues	${}^IE_t$
		All other credits taken	K_t'

Now the market equation for credits, including financial transactions, takes the form:

$${}^1S_t' + {}^DE_t + M_t^N = I_t' + {}^IE_t + {}^sE_t + {}^2M_{t+1} - {}^2M_t \qquad (D) \quad (8,1)$$

We may, if we prefer, combine the various more complicated models so far discussed into one single combined model. For this the market equation would run:

$${}^1S_t' + {}^DE_t + {}^2Y_t + N_t - {}^1I_t + M_t^N = _N^2I_t' + {}^IE_t + {}^sE_t + {}_R^2I_t + ({}^2M_{t+1} - {}^2M_t)$$
$$(A-D) \quad (8,2)$$

where the meaning of the new symbols will be clear from the preceding ones.

9. In the foregoing models it was assumed that there exists no *lag between income generated C_t+I_t and income paid out Y_t.* In reality however, there is such a lag; it takes some time for consumption and investment expenditures C_t and I_t to reach the ultimate producers. First, production takes place in various stages, and in the raw material stages, for example, the reaction of an increase in expenditures may be felt several months later than in the first stage, although this may be compensated by anticipating purchases of raw materials by the manufacturing stages. And second, some types of income, in particular dividends, are paid out several months after they are earned. Hence we have in the simplest case, where we assume that the lag is uniform for all industries:

$$Y_t = C_{t-\theta} + I_{t-\theta}. \tag{9, 1}$$

This implies that, in the case of model A, equation (5, 2) cannot simply be transformed into equation (5, 4) by omitting $Y_t-C_t-I_t$; instead we obtain:

$$K_t^n + (Y_{t+\theta} - Y_t) = K_t^E + I_t + ({}^2M_{t+1} - {}^2M_t) \tag{9, 2}$$

to be interpreted in the following way: part of the credits needed can now be supplied by the *"unintentional business hoarding"* $Y_{t+\theta} - Y_t$, accruing when at a period t income earned — which equals $Y_{t+\theta}$ — is higher than income paid out; or more credit is needed during a period when income earned is lower than income paid out. Similar changes can be made in the other models.

The lag just discussed works out in exactly the same way as the lag considered by Robertson and introduced by him as the lag between earned income and "disposable income"; his "disposable income" takes the place of our "income paid out." It is not quite clear, however, whether Robertson includes, in his lag, the elements just mentioned.

There is another very important lag in the relations to be studied — the "psychological lag" between changes in income paid out and the corresponding changes in consumption expenditures. This lag does not influence our *definition* of savings (4, 1), however, but only the way in which C_t, and hence also S_t', depend on income. In other words, it is a question of the determinants of S_t rather than of the definition. To this question we now turn.

10. In the preceding sections we have only *specified* the demand for a supply of credits; we have not yet gone into the question *what factors determine* them and what the relation between demand and supply and their determinants is. In the language of the ordinary market theory, we now want to describe the demand function, i.e. the functional relationship between (i) amount demanded and (ii) the rate of interest and the other "demand factors," which may also be named "data" or "determinants." Similarly for the *supply function*. It has become customary, particularly in the theory of interest, to replace this terminology by the well-known *ex-ante and ex-post terminology*. The subject of the present section may therefore also be formulated as follows: have we to interpret the symbols introduced in the preceding sections as ex-ante or ex-post concepts? Since it seems to me that several formulations of this terminology by the Swedish school are rather unsatisfactory, we may go into these questions at some detail now.

There is no difficulty about the meaning of ex-post concepts: they represent the realized values of the variables in question; and there is complete analogy between the market theory and the Swedish terminology: the realized values of "quantity demanded" and "quantity supplied" are equal. These realized values are, in essence, numbers and must be well distinguished from the demand and supply functions or schedules, a distinction which frequently is not made with sufficient emphasis. Its neglect has often led to unclear formulations. It is exactly the same question with the Swedish terminology. A good deal of clarification has been obtained by Ohlin's statement[5] that *expectations* or *ex-ante entities* are to be seen as *schedules;* in fact, they are nothing but the demand and supply schedules. This implies, first of all, that they depend on the yet unknown value of — in the case of the credit market — the rate of interest m. It also implies that the following two types of statement, very frequently used, are meaningless: (a) "ex-ante savings are unequal to ex-post savings"; and (b) "ex-ante savings are unequal to ex-ante investment." Statement (a) would mean that a schedule is compared with one special value of the dependent variable; can one say that the function $f(x)$ is unequal to the particular value that function assumes for $x = a$;

5. Cf. B. Ohlin, Alternative Theories of the Rate of Interest, Economic Journal 47 (1937), p. 423.

can one say, in other words, that $f(x) \neq f(a)$? Evidently not. Statement (*b*) implies that one schedule is unequal to another schedule, or that a function of some independent variables is unequal to another function of some other independent variables, which also has no meaning.

The confusion is, of course, due to the existence of a second meaning of the concept of ex-ante variables, viz. *one particular value of the schedule*. This particular value can only be obtained by assuming for the independent variables in the schedule some definite values. Taking, for example, some definite number for the interest rate m and for the other determinants, we get a definite value for savings ex-ante or anticipated savings. These values of both the independent and the dependent variables may be values that, with some right, are to be considered as the most probable values. But this second meaning of "expectations" must be carefully distinguished from the schedule meaning.

As already observed, the schedules will depend on other variables besides the rate of interest m_t. Some of them may be fixed beforehand; for example, credits granted in earlier periods, building plans, etc. Others, however, may be yet unknown variables; for example, income Y in the same period. In such cases the value of savings depends on more than one yet undetermined variable; it will itself only be determined at the moment m_t and Y_t will be determined by the trial-and-error process that leads to the market price m_t and, perhaps but not necessarily, simultaneously to Y_t.

The process of the formation of m_t and that of the formation of Y_t need not, in fact, be simultaneous. It is possible that m_t may be obtained by decisions based on some provisional value assumed for Y_t by the market parties — they may assume it to be equal to the realized value Y_{t-1} in the preceding period — and that, at the same time, Y_t is based on some provisional value for m_t, may be m_{t-1}. As already stated, however, we also may assume that both variables m_t and Y_t are formed simultaneously from the other data in the determining functions. The theory thus obtained is more approximative, is of a "coarser" type than the preceding one; the supposed adaptation or "trial-and-error" process by which the values m_t and Y_t are formed is neglected and passed over. It is the usual static method in economic theory that proceeds this way; it may be used for the general equilibrium of the whole economy as well as for partial equilibria.

But also the "finer" theory mentioned above, where m_t is assumed to be determined by the interaction of demand and supply — although not Y_t but Y_{t-1} appears as one of the determinants — does not consider the adaptation process in full detail. It is only in a *complete dynamic theory* that this happens. In such a theory an m_t found by trial and error does not fall from the air. The trials and errors themselves are described.

A simple example for a commodity market is the following. Let the quantity demanded at period t be x_t; it may depend on income Y_{t-1}, in the previous period, and the price p_t. Let the quantity supplied be y_t; it may depend on p_t and some other factor a_{t-1}. Now let p_t be announced at the beginning of period t; then, of course, there is no guarantee that $x_t = y_t$; in fact x_t will as a rule be $\neq y_t$. For a market where there exists, at every moment, a stock of unsold goods, this is no difficulty at all. We simply have that the stock w_{t+1} at the end of period t depends on the stock w_t at the beginning of the period by the equation $w_{t+1} = w_t + y_t - x_t$ and, as a rule, the value of the "announced price" p_{t+1} will depend on w_{t+1}: $p_{t+1} = f(w_{t+1})$.

In such a theory, where all successive steps are described separately, there is no variable that is dependent on a yet unknown value of another variable. Insofar as p_t in the course of some time approaches a certain limit, this limit is comparable to the outcome of the trial-and-error process in the "coarser" theories.

Only in such a completely dynamic analysis can all concepts be made exact. All other types of analysis raise the following difficulty. If the demand schedule, for example, depends, inter alia, on the anticipated value, say of Y_t, then at the beginning of the trial-and-error process this value may be another than during or towards the end of the process when it approaches the ex-post value of Y_t. Now m_t is the product of this process; on what anticipated value of Y_t does it depend? Does it depend on the anticipated value of Y_t at the beginning of the period t, which more or less rightly may be considered as "the" ex-ante value of Y_t, or does it depend on the "ex-post" value? This difficulty only fails to arise when we consider the successive steps of the adaptation process.

There is not one ex-ante value (in the second sense) of a given variable; there is a whole series, depending on the time point to which the expectation relates; and they successively approach the

ex-post value. For small intervals, therefore, the difference between ex-ante and ex-post values is not large. Myrdal[6] however, is right in contending that by making the intervals smaller the difference cannot be made to vanish altogether. What matters, however, is on what determinants the expected values or the schedules are based. And this leads us back to our main subject.

11. We start our discussion of the *determinants* by considering the determinants of active *savings* S_t. The chief determinant is certainly the size and distribution of the national income Y_t; we need not add very much to what Keynes and others have taught us on this subject.[7] One aspect, however, deserves a closer examination — the question of the *psychological lag* that may exist between a change in income and a corresponding change in savings. An increase in income does not immediately lead to a change in consumption or saving habits; and the possibility exists that savings as well as consumption expenditure lag behind incomes. The simplest case is present if this lag is not a distributed, but a simple one; i.e. that $S_t = f(Y_{t,\eta})$, where η is the lag. Assuming for a moment that consumption expenditure C_t also depends on $Y_{t-\eta}$, we find that during the elementary period t the total amount handed over to others, viz. $S_t + C_t$, is also a function of $Y_{t-\eta}$. This amount may or may not be equal to $Y_{t,\eta}$. If it equals $Y_{t-\eta}$, we shall say that there is no intentional hoarding; if it does not equal $Y_{t-\eta}$, we shall call the residual *intentional (family) hoarding* ^1H_t; hence

$$^1H_t = Y_{t,\eta} - C_t(Y_{t-\eta}) - S_t(Y_{t-\eta}). \tag{11, 1}$$

Apart from intentional hoarding there will appear, as a consequence of the lag η, an amount of *unintentional (family) hoarding:*

$$^1H_t' = Y_t - Y_{t-\eta} \tag{11, 2}$$

which does not occur when $\eta = 0$.

The lag we now discuss is not of a technical nature. In some respects, however, a technical lag, due to the organization of banking, will work out in a similar way. In a sense one might say that income Y_t only becomes "disposable" (namely, in the imagination of the subject) at the period $t+\eta$ and is then distributed

6. G. Myrdal, Monetary Equilibrium, London, 1939, p. 61.
7. E. A. Radice (Savings in Great Britain, 1922–1935, Oxford, 1939) has shown that business savings also are intimately connected with income, i.e. profits.

between consumption, savings-handed-over and intentional hoarding. The income difference $Y_t - Y_{t-\eta}$ accrued in the meantime represents unintentional hoarding; in the case of the psychological lag this hoarding is located in the cash of the income earner; in the case of a technical lag it will take place in the cash of some business firm, or perhaps of a bank.

It is possible to define our types of hoarding also in the case of a distributed lag between Y and S; it would take us too far afield, however, to go into this matter.

Apart from the question of the lag, there is that of the shape of the functional relation f between S_t and $Y_{t-\eta}$. This relation may be called, following Keynes to some extent, the propensity-to-save function. One of the simplest cases possible is that of a linear function:

$$S_t = \sigma_0 + \sigma_1 Y_{t-\eta} \qquad (11, 3)$$

where σ_0 and σ_1 are constants. Assuming this shape does not, as is frequently thought, imply proportionality between income and savings; proportionality would require σ_0 to be zero. The rather meagre statistical material at our disposal is not violated very much if we take $(11, 3)$ as the propensity function; although some authors[8] suggest a constant elasticity (i.e. exponential) function. The advantage of formula $(11, 3)$ is that σ_1 at once represents the marginal propensity to save, i.e. the increase in savings corresponding with a (small) unit increase in income.

A more complicated representation of the propensity function could take account of a possible and probable difference between long-term and short-term marginal propensities; this would be possible if we put

$$S_t' = \sigma_0 + \sigma_1 Y_{t-\eta} + \sigma_2 \bar{Y}_{t-\eta} \qquad (11, 4)$$

where $\bar{Y}_{t-\eta}$ is an average of previous Y-figures, say those prevailing in the ten-year period before $t - \eta$.

Apart from this possibility, there is the one that σ_0 and σ_1 are not exact constants, but themselves dependent on other variables; in fact $(11, 4)$ is a special case of this state of affairs. There is not very much known about all this; as a curiosity I mention King's figures on incomes and expenditures of the various classes

8. E. Gilboy, Review of Economic Statistics, 1939.

of society in the seventeenth century,[9] from which a propensity to save can be deduced not very different from the present figure.

Savings may also depend on *other variables* than income, the second variable to be thought of being, of course, the rate of interest. We need not repeat the well-known discussions on this topic and will not encounter much opposition when stating that a very clear dependency of total savings on the rate of interest has never been demonstrated: positive and negative influences seem to balance pretty well. It is quite another question that some particular types of saving do depend on the rates of interest connected with these types; this is a question of substitution elasticity rather than total elasticity.

It goes without saying that there are many other factors that influence savings, such as the opportunities offered, the general level of thriftiness of a population or of an epoch; so very little exact knowledge of them can be presented, however, that we abstain from discussing them here.

12. We proceed to a discussion of the determinants of our next supply item, *depreciation allowances*. An important general remark about this item is that it is very insensitive to cyclic changes. This is a consequence of its depending on the investments made during a long period — in fact, all investments not yet written off at the time period in consideration. Another reason why there are no pronounced cyclical fluctuations in depreciation allowances may be the rather conventional way of their calculation.

Their long-run movements will depend on the general growth in the value of production and the changes in capitalization and in durability of capital goods. An increasing amount of capital per unit of product will require, other things being equal, an increase in depreciation allowances per unit of product. An increase in the durability of capital goods would cause a fall in depreciation allowances. Colin Clark states that there is such an increase in depreciation charges that we must assume a considerable decrease in the durability of capital goods for the last few decades.[1]

9. Cf. Colin Clark, National Income and Outlay, p. 212. Since this text was written, important contributions to our knowledge on this subject have been published, i.a. by Woytinski, Bean and Klein.

1. C. Clark, National Income and Outlay, London 1937, p. 184.

13. The determinants of *credit creation* are to be sought in the sphere of banking policy. For long periods, the changes in the central bank's gold stock will have been outstanding among these determinants. This is affirmed by the figures given by Warren and Pearson, showing a close parallelism — in long-run movements — between gold stock and total money in circulation and between the long waves of the gold stock and those of the price level, already pointed to by Cassel. This parallelism ceased to exist, however, after 1914. Since then, the newer ideas on banking policy as well as the necessities of warfare have given to credit creation a still more exogenous character.

Also in the short run, the fluctuations often diverge. Many other factors intervene; as one of the most evident and most important examples, we again quote the financing of wars and the provisions in periods of trade crisis. The factors then governing the creation of credit are often of an incidental character, and a general description would be very difficult.

The rate of interest itself also exerts an important influence, particularly on that part of credit creation due to commercial banks. This seems to be well illustrated by the results of some correlation calculations presented in my investigation on American business cycles between 1919 and 1932.[2] There it was found that both the "supply of deposits" and the "demand for short claims" by commercial banks showed a clear dependence on the short-term interest rate.[3]

For the credit creation by *central banks* the causal connection — in the short run, at least — just works the other way round: it is the discount rate that is determined autonomically by these banks, and the amount of credit created is determined by the demand for it corresponding with the discount rate.

14. *The purchases of securities*, a further supply item in the credit market, as well as the *sales of "old" securities*, will first of all depend on their price level, which, in connection with their nominal yield, determines their real yield, which is nothing but one or another type of interest rate. Important determinants will be the expected price movements and, in the case of stocks, the

2. J. Tinbergen, Business Cycles in the United States of America, 1919–1932 (Statistical Testing of Business Cycle Theories, II), League of Nations, Geneva, 1939.

3. A fuller discussion of the exact meaning of these results is presented in section 24 below.

expected dividend. Expected dividends are often based on the last-known dividends, but insiders will know more about the actual state of affairs and outsiders will try to judge it from the general course of business. As to the anticipated prices, two ways of forming a judgment must be clearly distinguished. Experts will base their judgment as much as possible on objective criteria, leading them to some level of the price they consider appropriate, and they will assume that sooner or later this level will be realized. They will base their speculations on this type of consideration. The "great public," when coming into the market, will, however, often follow another way of speculating. They will extrapolate past movements more or less automatically and will be impressed by price gains made some time before. Their purchases will, after having passed a certain level, lead to "booms" of a rather dangerous character, where prices lose contact with reality, and the reaction may be a crisis. A statistical illustration of this type of reaction is found in the League of Nations investigation already referred to (cf. Section 4, 8: "The share price equation," page 106).

15. Turning to the demand side, we first discuss the determinants of *new investment*. We shall not do this in great detail, since we paid a good deal of attention to this matter in another League of Nations investigation.[4] Hence we only summarize the results obtained there, adding a few remarks where it seems useful.

New investment may be split up into new investment in *durable goods* and new investment in *non-durable goods* (stocks of raw materials, semi-manufactures and products: working capital).

Investment in *durable goods* is — in the short run — primarily determined by the margin between anticipated profit and the interest rate. Since profit rates (measured as a percentage of capital invested) show much wider fluctuations than interest rates, the latter have a less pronounced influence than the former. Profit anticipations seem to a great extent to be determined by real profits made some time before; in addition, incidental factors such as new inventions or the opening of new markets will have an influence, but it is difficult to detect this influence by means of statistical investigations, and it is at any rate remarkable how

4. J. Tinbergen, A Method and its Application to Investment Activity (Statistical Testing of Business Cycle Theories, I), League of Nations, Geneva, 1939.

close a correlation exists between investment activity, on the one hand, and a combination of factors among which profits half a year or a year before stand out, on the other hand.[5] Presumably, the influence of inventions and new markets are partly long-run influences, explaining why there are certain epochs of high and others of low investment activity, and manifesting themselves in the trend of investment; partly so-called incidental or irregular influences, showing themselves in the "residuals."

For some particular types of investment, other determinants were found to exist; those mentioned by the acceleration principle, for example, in the case of investment in rolling railway stock. It is remarkable, however, that only very faint traces of the action of this principle could be found in general investment activity.[6]

Investment in *non-durable goods* presents a particularly difficult question: it is difficult to disentangle and to summarize by a general formula the numerous factors influencing speculation in *raw materials*. Here our knowledge is very restricted. As to stocks of *finished products*, the chief impression is that there the acceleration principle works very clearly.[7] There is a pronounced tendency for these stocks to move parallel — with some lag — to general activity; this implies that the increase or decrease in stocks — i.e. net investment — moves parallel to the rate of increase or decrease of general activity. Possibly, also, the rate of interest exerts an influence, but this has not yet been very clearly established by statistical tests.[8] It is worth noting that on this point the theorists themselves mentioning this influence are not always clear in their formulations: is it total stocks or is it investment in stocks — hence the rate of increase — that is supposed to be governed by the rate of interest?

16. *Reinvestments* in durable goods are, first of all, subject to the so-called echo-principle, stating that high reinvestment activity is to be expected one lifetime period after a top of high new investment in durable goods. If the average lifetime of a

---

5. Cf. particularly graphs III 2 (p. 56), III 3 (p. 57), III 4 (p. 58), III 5 (p. 59), III 9 (p. 81), III 10 (p. 82) and III 11 (p. 83) of the last-mentioned investigation.

6. Cf. Section 10 and Chapters IV and V.

7. Cf. J. Tinbergen, Business Cycles in the United States 1919–1932, section (2–6), p. 49.

8. The influence found in the cited investigation is rather weak and hence dubious.

ship is taken as seventeen years, then seventeen years after a shipbuilding boom an "echo" of this boom will be found in replacement. The working of this echo is soon dissipated, since there is a considerable spread in lifetime among individual capital goods as well as among various categories. It has even often been doubted whether there is any evidence at all of its working; I think that such evidence has been presented in the case of shipbuilding[9] and of automobile production.[1]

There are, however, important further determinants to reinvestment activity. Outstanding seems to be the influence of the general business situation: there is a clear tendency to restrict reinvestment in depression periods and to accelerate it in booms. This is clearly shown by the just cited investigations.

As to the long-run determinants, these are largely the same as those for depreciation allowances, since in the long run these two variables will tend to be equal. Hence we may refer to section 12.

17. Finally we have to consider the determinants of *cash holdings* $^2M_{t+1}$. It is not necessary to indicate those of ^2M_t, since initial cash holdings have to be considered as given for each unit-period; they are the result of past actions. The desired cash holdings $^2M_{t+1}$ are, generally speaking, the result of weighing off against each other the advantages of a greater cash and the disadvantages of paying more interest if more credits are to be held. Before going into the motives for holding cash we may discuss some implications of this quite general statement. Indicating by M the desired cash, by Y^I the interest payments, and by $\Omega(M, Y^I)$ the utility function to some individual, this individual will try to maximize Ω. Changes in M and Y^I are, however, not independent. Interest charges are the total of all interest due on the various credits outstanding; a credit $_TK_t^o$ outstanding at the beginning of period t, and originating from the period T bears a charge of $_TK_t^o m_T$, where m_T is the interest rate that existed at the period T. Total interest charges are, therefore, equal to

$$Y^I = \Sigma T_T K_t^o m_T. \qquad (17, 1)$$

9. Cf. J. Meuldijk, Der englische Schiffbau während der Periode 1870–1912 und das Problem des Ersatzbaues, Weltwirtschaftliches Archiv 52 (1940), p. 524; J. Einarsen, Reinvestment Cycles, Oslo, 1938.

1. The Dynamics of Automobile Demand, edited by the General Motors Corporation, New York, 1939; P. de Wolff, De vraag naar personenauto's in de Vereenigde Staten, De Nederlandsche Conjunctuur, November 1936, p. 18.

An extension of cash holdings M at the period t can be obtained by an equal extension of credits $_tK_t^o$, which bear an interest charge of m_t, interest rate at period t. The maximum value for Ω is obtained for such a value of M that

$$\frac{d\Omega}{dM} = \frac{\partial\Omega}{M} + \frac{\partial\Omega}{\partial_tK_t^o} = \frac{\partial\Omega}{\partial M} + \frac{\partial\Omega}{\partial Y^I} \; m_t = 0. \qquad (17, 2)$$

This equation expresses in the well-known way that the increase in utility caused by the increase in M just outweighs the decrease as a consequence of the increased interest charges. It may be solved — in principle, at least — for M and then yields a relation

$$M = F(m_t \ldots) \qquad (17, 3)$$

between M and m_t, which in a sense is Keynes' liquidity preference function (but then related to total cash holdings, not to "inactive balances" only).[2]

One point we want to emphasize is that this function F depends not only on the rate of interest m_t during the period under consideration, but *also on a number of earlier interest rates*, determining the interest charges on older credits. Another point is that it is not always possible to adapt M to one's desires simply by changes in credits outstanding at the rate of m_t. From the moment where all recent credits outstanding are redeemed and a further desire for a diminishing of cash holdings would exist, the individual under discussion has to redeem "older" credits — if possible — and this means that another, viz. an "older" interest rate m_T comes into play. This implies that M_{t+1} cannot, under all circumstances, be considered as an invariant function of m_t alone; which seems to be a clear departure from the Keynesian doctrine. One of two alternatives comes instead: either older m_T's come into the liquidity preference function $(17, 3)$, or if no opportunities for further redemptions exist — they may be forbidden by the older contracts — a value of M deviating from that in $(17, 3)$ will be chosen by the individual.

The further determinants appearing in $(17, 3)$ will be those so well described by Keynes; they are the incorporations of the transactions, precautionary and speculative motives and the desire for investment finance. The transactions motive, for example, leads to total transactions as a determinant; further,

2. Cf. G. Haberler, *Prosperity and Depression*, 3d edition, Geneva, 1941, p. 210.

such factors as the rate of increase of the price level of both goods and securities may come in, or the fear for illiquidity as a consequence of a crisis. This latter factor will be of particular importance shortly after a crisis.

18. A few words may be said about the *length of the elementary period*. As a minimum this period should be equal to the period for which the rate of interest in consideration is fixed; for most rates this will be one day. It does not seem necessary that during this period plans are not changed, as Ohlin claims.[3] The process of the formation of the interest rate is a process of trial and error, as we have seen, if we use a semi-static analysis (cf. section 10 above) and during this process plans may be changed.

Taking a day for the elementary period means making a very detailed analysis. For most practical investigations it will be convenient — and, in our opinion, quite acceptable — to take longer periods of reference, such as months, years, or even decades; this last period, if one is interested in long-term movements of interest rates. There is another reason why taking such longer periods may be useful. As Lerner[4] rightly observes, a certain minimum length is required in order to permit all subjects involved to adapt their holdings of credits and cash to the interest rate; if we want to assume — for reasons of simplification — that such an adaptation takes place, we are obliged to take such a longer period. Or, to put it the other way round: if we take a very short period of reference, such as a day, we must, in accordance with what was said in section 17, admit that in the liquidity preference function other, "older" interest rates play an important rôle and that the Keynesian approximation with only the present rate is inadequate.

It is perhaps superfluous, but at any rate useful, expressly to state that the period of reference now discussed is not necessarily the same as one of the lags discussed in sections 9 and 11; in particular, Lutz[5] has made a number of interesting points on the implications of this fact.

Neither is the elementary period the same as the period over which the financial plans are made, referred to above. For the

3. B. Ohlin, Some notes on the Stockholm Theory of Savings and Investment, Economic Journal XLVII (1937), p. 61.

4. A. P. Lerner, Alternative Formulations of the Theory of Interest, Economic Journal XLVIII (1938), p. 211.

5. F. A. Lutz, The Outcome of the Saving-Investment Discussion, this JOURNAL, LII (1938), p. 588.

flotation of a bond issue, for example, a plan will be made referring to some decades, dependent, perhaps, on the lifetime of some capital goods involved or on the presumable period of peace in which a war loan is redeemable. Particularly for the bond market it is clear that if we take a very short elementary period — to be justified by the fact that a quotation of bonds is made up every day — the chance is very great that not all holders of these securities will, in each elementary period, adapt their holdings to the state of the market.

If we take longer periods of reference we do get, of course, a more approximative result. It is equivalent to neglecting a number of links in the process of adaptation; it implies a type of analysis nearer to static theory (cf. section 10). For some purposes, particularly for the explanation of the long-run trend of interest rates, it will be very useful, however. Taking a long period of reference has one very important consequence: the influence of the flow items in our market equation (5, 6; 6, 6; 7, 3 or 8, 1) becomes more and more preponderant as compared with that of the stock items, and hence the influence of the liquidity preference schedule diminishes in favor of that of savings and investment determinants.

19. Before discussing (in Part II, below) some applications of our market formula, let us consider for a moment the *controversial formulations* by Keynes and what he calls classical authors. The essence of the Keynesian treatment of the problem of determining the rate of interest may thus be summarized:

(*a*) The rate of interest cannot be determined by an equation: savings = investment, where savings represents the supply of, and investment represents the demand for investible funds; the impossibility of doing so is a consequence of the fact that the adaptation of savings to investment is obtained by income changes rather than changes in interest rates. In mathematical language, not the rate of interest m, but (national) income Y is the independent variable in the equation savings = investment.

(*b*) Hence, the determinants of savings and investment do not influence m, since for any value of m there is equality of savings and investment.

(*c*) The rate of interest, on the other hand, is determined by the liquidity preference function, connecting the quantity of money with m.

(*d*) This implies that the rate of interest must be understood

by a stock analysis, viz. an analysis of the stock of money present at a given moment, rather than by an analysis of the flows of savings and investment per unit period.

There is some advantage in commenting on these statements with the help of a concrete case, and we choose our simplest model A (equation 5, 6) as such a case. The first point to be emphasized is that *Keynes' "Savings" are not ours;* indicating his by S_t^K, we have:

$$S_t^K = S_t + {}^1H_t \tag{19, 1}$$

where ${}^1H_t$ represents total hoarding (intentional ${}^1H_t'$ and unintentional ${}^1H_t''$) by families:

$$^1H_t = {}^1M_{t+1} - {}^1M_t = {}^1H_t' + {}^1H_t'' \tag{19, 2}$$

The second point is, that *the equations: "demand for investible funds equals supply of such funds" and "Keynesian savings equal investment" are not identical;* the first one runs, in our symbols:

$$S_t + M_t^N + {}^2M_t = I_t + {}^2M_{t+1} \tag{19, 3}$$

whereas the second is:

$$S_t^K = S_t + {}^1H_t = I_t. \tag{19, 4}$$

Nor is there a contradiction between these two equations: Deducting (19, 4) from (19, 3) we obtain:

$$M_t^N + {}^2M_t = H_t + {}^2M_{t-1} \tag{19, 5}$$

which may be written in either of two forms:

$$M_t^N + {}^2M_t + {}^1M_t = {}^2M_{t+} + {}^1M_{t+1} \quad \text{(``stock analysis'')} \tag{19, 6}$$

$$M_t^N = {}^1H_t + {}^2M_{t+1} - {}^2M_t \quad \text{(``flow analysis'')} \tag{19, 7}$$

Of these two equations, the first expresses that the amount of money made available by the banks (money already present at beginning of period: ${}^1M_t + {}^2M_t$, plus new creation of money M_t^N during period) equals the demand for money ${}^1M_{t+1} + {}^2M_{t+1}$ at the end of the period, determined by the liquidity preference functions of families and firms. The second expresses that the newly created money equals the net flow of hoarding by families and business firms. As indicated in brackets, the first equation takes the form of a stock analysis, the second of a (net) flow analysis. As Ohlin has rightly pointed out,[6] these two ways of representing the situation (he calls them "gross" and "net" formulation) are equivalent. In fact, their relation is the same as between a function and its differential coefficient (more exactly: its first dif-

6. B. Ohlin, loc. cit., in particular pages 224 and 225.

ference) or between the integral (more exactly: the cumulation) of a function and that function.

Nevertheless, the whole controversy between stock and flow analysis is not solved with this statement. There are two further questions involved. The first is that, dependent on the determinants (or, if one likes, the independent variables) behind these entities, *one of the two formulations may be a better instrument of analysis than the other.* If it were true that the total stock of money present at any moment is adapted immediately to the then prevailing rate of interest and $^2M_{t+1}+{}^1M_{t+1}$, were an invariant function of that rate, the stock analysis would be a better instrument. For if the stock $^2M_{t+1}$, for example, is a function of m_t only, then the flow $^2M_{t+1}-{}^2M_t$ is a function of the two interest rates m_t and m_{t-1}, and hence a more complicated expression. But if only the flow of new money is dependent on that rate, whereas the older "parts" of the stock are dependent on "older" rates, the flow analysis is the more appropriate instrument.

For now the stock would be a more — and even a far more — complicated expression in terms of the interest rate: it would, in principle, depend on all previous interest rates. Thus there remains the question to be answered, which of the two — or even more, namely, "mixed" — alternatives is closest to reality. Lerner rightly argues that the flow analysis in its simplest form cannot be right: a certain interest rate cannot induce people continuously to hoard a certain amount of money per time unit. On the other hand, our analysis in section 17 seems to prove that the pure stock analysis — at least, if applied to very short periods — is not correct either. It is only so if we do not consider too short periods; it is an approximation.

This brings us into the second question. It is useful to make a distinction between *two sorts of flow analysis, net flow and gross flow analysis.* The difference can best be clarified by reminding the reader of our concepts (cf. section 2) K_t^n and K_t^N, where K_t^n was the total amount of new credits accorded during time period t and K_t^N was the net increase in credits outstanding; we call K_t^n the gross flow and K_t^N the net flow of new credits. (Of course, similar distinctions may be made for cash holdings). The connection between K_t^n and K_t^N is given by equation $(2, 4)$: $K_t^N = K_t^n - K_t^E$, where K_t^E represents credits expired. The difference in behavior between K_t^N and K_t^n is best illustrated by the

consideration of two special cases. First, assume that all credit contracts are accorded for one elementary time period only. That implies that credits outstanding at the end of period 1 equal credits accorded during that period:

$$K_2^o = K_1^n \qquad (19, 8)$$

and credits expiring during period 1 equal credits outstanding the beginning of period 1:

$$K_1^E = K_1^o. \qquad (19, 9)$$

Equation (19, 8) expresses that now *stock = gross flow;* for the net flow, however, we still have that it equals the rate of increase in stock.

Secondly, assume that all credits are accorded for two elementary time periods. This means that, instead of (19, 9), we now get:

$$K_1^E + K_2^E = K_1^o \qquad (19, 10)$$

telling that credits outstanding at the beginning of period 1 will now be expiring during the next two periods. Now since

$$K_3^o = K_1^o + K_1^N + K_2^N = K_1^o + (K_1^n - K_1^E) + (K_2^n - K_2^E) \qquad (19, 11)$$

it follows that:

$$K_3^o = K_1^n + K_2^n \qquad (19, 12)$$

or: *the stock equals the gross flow taken over two consecutive periods.* Still the net flow, if taken over these two periods, equals the rate of increase in the stock. The stock therefore equals the cumulation of the gross flow over a *finite* time interval, and the cumulation of the net flow over an *infinite* time interval. The relations get more complicated if the contracts are of unequal length, and we cannot go into this here. But it will be clear that there is conceivable a third type of theory, namely, one where the simplest description of reality is obtained by expressing the *gross* flow by some function of the independent variables. From the preceding sections it follows that we consider this the best solution of the stock-and-flow antithesis.

To state clearly what determinants will appear in the system of equations, we rewrite the two equations, in both the "classical" and the Keynesian form, using function symbols where that seems appropriate.

"Classical":

$$S_t(Y_t, Y_{t-1}, m_t) + M_t^N(m_t) + {}^2M_t = I_t(Y_{t-1}, m_t) + {}^2M_{t+1}(Y_t, m_t)$$

<div align="right">(C 1)</div>

$$Y_t = C_t(Y_t, Y_{t-1}) + I_t(Y_{t-1}, m_t)$$

<div align="right">(C 2)</div>

Keynesian:

$$M_t^N(m_t) + {}^2M_t + {}^1M_t = {}^2M_{t+1}(Y_t, m_t) + {}^1M_{t+1}(Y_t, m_t)$$

<div align="right">(K 1)</div>

$$Y_t = C_t(Y_t, Y_{t-1}) + I_t(Y_{t-1}, m_t).$$

<div align="right">(K 2)</div>

In addition, the *identity* connecting C, S and ${}^1M$ (our equation (4.1)) may be repeated:

$$Y_t = C_t + S_t + {}^1M_{t+1} - {}^1M_t.$$

<div align="right">(4.1)</div>

This, however, is indeed an *identity*, meaning that it must be true, for any value of Y_t and m_t, whereas the above *equations* are conditions to be fulfilled by Y_t and m_t.

Generally speaking, the pair of equations (C 1) and (C 2), or the equivalent pair (K 1) and (K 2), determines simultaneously the variables Y_t and m_t. The determinants occurring in these equations are those of S, I, ${}^2M_{t+1}$, C, or, since there is a definitional connection (4. 1) between S, C and ${}^1M$, the determinants of S, I, ${}^2M_{t+1}$ and ${}^1M_{t+1}$. The latter two taken together forming the liquidity preference (of firms and families, respectively) it may be concluded that, generally speaking, *the determinants of (active) savings, investments and liquidity preference together determine Y_t as well as m_t*. Only in the special case, where, in equation (19.6), ${}^2M_{t+1}$ as well as ${}^1M_{t+1}$ would only depend on m_t and not on Y_t, Keynes' thesis that m_t is determined by liquidity preference only, would hold. Since Y_t generally will influence ${}^1M_{t+1}$ and ${}^2M_{t+1}$ through the transactions motive, this special case does not seem to be of great significance.

On the other hand it does not seem probable that in S_t or I_t the variable m_t does not occur; it is probable that, at least in I_t, m_t appears as an independent variable; Y_t is not, if this is so, the only independent variable determining the amounts S_t or even S_t^K.

Here we are led to exactly the same conclusion as Hicks[7] and Pallander[8] — that Keynes' theory is only a special case of a more

7. J. R. Hicks, Mr. Keynes and the Classics, Econometrica V (1937), p. 147.

8. T. Pallander, Keynes' allmänna teori och dess tillämpning inom ränte-, multiplikator- och pristeorien, Ekon. Tidskrift XLIV (1942), p. 233.

general one. This seems to have been overlooked by Lerner[9] and perhaps also by Haberler, when these authors try to show that there is, in the end, no difference of opinion about facts, only about definitions, between Keynes and the "classics."

II

20. So far we have treated the credit market as a single market. It is, however, composed of various compartments, more or less isolated, with different prices. The method of dealing with this phenomenon of partial markets within the framework of the foregoing analysis may be illustrated by considering in more detail the roughest distinction between various types of credits we are accustomed to work with, namely, that between "long" and "short" credits, to be indicated by left-hand bottom indices L and S. To fix the ideas, the border line between long and short credits may be taken at one year. There are two representative interest rates $_Lm$ and $_Sm$; say the yield on bonds and the private discount rate. Each item in the market equation has now to be split up into an item for long and one for short credits; the demand item $_NI$, e.g. into $_{LN}I$ and $_{SN}I$, where $_{LN}I$ may be the demand for bond (and stock) capital to finance new investments and $_{SN}I$ the demand for bank credit for the same purpose ("anticipation credits"). Both will depend on the determinants relevant to new investments and, in addition, on both interest rates. In particular, it will be the relation of these two rates and the expected future long-term rate and its movements that will influence the method of financing the investments. Attempts to proceed to econometric analysis in this way have been made in my above-quoted investigation on the business cycles in the United States 1919–1932, to be discussed further in section 26.

Applying the same idea in the other items of the market equation, we get two equations, expressing the equality of demand and supply respectively for short and for long credits:

$$_S^D K^n = _S^S K^n \qquad (20, 1)$$

$$_L^D K^n = _L^S K^n. \qquad (20, 2)$$

9. Loc. cit. Unfortunately Lerner's interesting paper is somewhat obscured by two circumstances. First, he is not clear about what definition for savings he assumes, and secondly, he uses the symbol Y alternatively for income and for "ordinate" generally; thus he also uses Y in one or two graphs to mean the interest rate, which, in view of the subject under discussion, is somewhat unfortunate.

In order to exemplify the results that may be obtained by our instrument of analysis, let us go into somewhat greater detail.

As our special purpose for the following few sections we take the explanation of the cyclical behavior of the two interest rates. Hence we shall neglect all determinants that are only of importance for the explanation of trend movements or of very short fluctuations: only determinants showing an important and systematic cyclic component will be considered. The cyclical movement of the non-credit markets will be considered as given and will be assumed to be represented by the fluctuations in the general profit rate z, since most of the other economic variables show a cyclical pattern parallel to that of z.

Demand for short credits will chiefly be for new investment purposes, for speculation and for increasing liquidity; new investment only as far as it takes the form of investment in commodity stocks or as the credit takes the form of an anticipation credit. The supply of short credits originates from the banks. Demand for long credits will be for investment in durable goods: capital issues by public authorities, public utilities and mortgages for building or for some types of consumption credit (public authorities and the general public for as far as this demand materializes by the selling of "old securities"). The supply of long credits originates from the savings of family households and from the investment of part of the banks' resources. The concrete shape of the relation between each of these items and the two interest rates and the other determinants will depend on (i) the planning period or "horizon" of the subjects involved; (ii) the time shape of the need for or the availability of credit within that time period, determining, for example, the degree of idleness of long credits and (iii) the technical opportunities of demanding or supplying (the question whether one is or is not permitted to enter a certain market) and, closely related, the possibilities of substitution.

If, in particular, some type of investment promises a profit rate of z, whereas the interest rate to be paid is m, demand will depend on $z - m$; if, in addition, there is a possibility of lending in another market compartment at a rate m', it will also depend on $m' - m$.

For part of the demand for short credits, namely, commercial and speculative credits, no substitution possibilities exist; they will depend on the difference between the general profit rate z

and $_Sm$ only. For advances, the substitution by a long credit exists, and hence they will, in addition, depend on $_Lm-_Sm$. We are led to a somewhat more exact theory by the following considerations. The choice may be said to be between:

(a) to continue a short credit for a period θ and then to take a long credit;

(b) to start the long credit immediately.

The first possibility implies the payment of a rate of $_Sm$ over a period θ, instead of $_Lm$, and a payment of another long-term interest rate than the one now prevailing, namely, the rate expected to exist after a lapse θ. If the investor bases his expectation on an extrapolation of the present movement of $_Lm$, the expected rate will be $_Lm+\theta_L\dot{m}$. Thus interest payments over a horizon T will be, in case (a) $\theta_Sm+(T-\theta)$ $(_Lm+\theta_L\dot{m})$ and in case (b) T_Lm. The choice will depend on the difference: θ $(_Sm-_Lm+(T-\theta)_L\dot{m})$. From this somewhat closer examination we learn that, apart from the difference $_Sm-_Lm$ already mentioned, the choice will depend on $_L\dot{m}$, if the expectations are based on extrapolation of the present movement. If the expression θ $(_Sm-_Lm+(T-\theta)_L\dot{m})$ increases there will be a stronger demand for long credits; if it decreases, the demand for short credits will be stronger. In view of the foregoing considerations, demand for short credits may be assumed to be given by the equation

$$_S^DK^n = a_1(z-_Sm) - a_2(_Sm-_Lm) - a_{3L}.\dot{m} \qquad (20, 3)$$

The *supply of short credits* originates, as we stated already, from the banks. Disregarding incidental factors such as fluctuations in gold stock, it will primarily depend on the short interest rate itself, which determines the degree of profitableness of creating credit. These credits may, however, partly be used for buying bonds; this again depends on the market situation and prospects. A short credit, say, for a time period θ, will yield the bank interest receipts proportional to $_Sm\theta$, whereas an investment in bonds yields (i) interest receipts proportional to $_Lm\theta$ and (ii) a price gain or loss proportional to p, the difference between the expected price of the bonds at time $t+\theta$ and the price at time t. Assuming again that the expected price movement is based on extrapolation of the existing movements, it may be deduced from $_Lm$ and $_L\dot{m}$ in the following way. Let the initial bond price be 1, then the price

at time $t+\theta$ will be

$$\frac{_Lm}{_Lm+_L\dot{m}\theta},$$

the price gain will be

$$\frac{_Lm}{_Lm+_L\dot{m}\theta}-1=-\frac{_L\dot{m}\theta}{_Lm}$$

approximately. Together with interest receipts this shows a difference with interest receipts in the short credit case of

$$(_Sm-_Lm+\frac{_L\dot{m}}{_Lm})\theta$$

Again the supply of short credits will therefore be the greater, the higher this difference is. Taking account of this substitution effect as well as with the absolute profits to be obtained by credit creation, the supply of short credits will be, in the simplest case:

$$_S^SK^n=\beta_{1S}m+\beta_2(_Sm-_Lm)+\beta_{3L}\dot{m}. \qquad (20,4)$$

The *demand for long credits* will chiefly be governed by the difference between profit expectations, which we assume to be dependent on the actual profit rate z, and the long-term interest rate $_Lm$. For another part, e.g. as far as it is demanded by public authorities for covering deficits and for consumption purposes, it will depend on national income and even in a negative way; but since the fluctuations in national income are closely correlated with those in z, this only means adding a term of a type already introduced. It may be, however, that this part of demand is not so much influenced by the interest rate itself, and hence the term with $_Lm$ need not have numerically the same coefficient as the one with z. In addition, also here, a substitution effect — the complement to the effect discussed for the demand for short credits — is to be expected. It will be based on the same variables as the latter, but will show the opposite sign. Hence the demand for long credits will be of the following shape:

$$_L^DK^n=\gamma_1z+\gamma_2(_Sm-_Lm)+\gamma_{3L}\dot{m}-\gamma_{4L}m. \qquad (20,5)$$

Finally, the *supply of long credits* originates from two sources, private and banking investments. Private investments may be out of current savings, which largely depend on national income and hence approximately on z, and out of accumulated wealth from older savings, held in money form. The latter may be assumed

to depend on $\int_0^t z\,dt$, the accumulated value of previous profit rates — because of the close correlation between profit rates and national income. They may depend also on the long-term rate of interest, but do not seem to do so in a very marked way. Investment by banks will be much more influenced both by that rate of interest and by the substitution motive, which again is the counterpart of the substitution effect in the supply of short credits. In addition it depends on liquidity considerations which will, however, largely be determined by the general business situation, i.e. by z. Hence we may put:

$$\tfrac{S}{L}K^n = \delta_1 z - \delta_2({}_S m - {}_L m) - \delta_{3L}\dot m + \delta_{4L}m + \delta_5 \int_0^t z\,dt. \qquad (20, 6)$$

One interesting implication of the above analysis may be emphasized beforehand — the rôle attributed to savings. If it is true that family savings are chiefly led to the long-term credit market, two conclusions may be drawn: (i) they will not influence in a direct way the short interest rate, and (ii) even their influence on the long-term interest rate will be very smooth, because this latter is chiefly governed by the accumulated fund of anterior and present savings: a considerable change in annual savings only means a slight change in this fund. Abundance of savings will not therefore result in a sudden change in interest rates; shortage will only do so if the banking system is operating exactly at the limit of its reserves and cannot even supply the smallest extra amount of complementary means of payment. These conclusions seem rather interesting. The specifications discussed in this section transform the equations (20, 1) and (20, 2) into:

$$a_1(z - {}_S m) - a_2({}_S m - {}_L m) - a_{3L}\dot m = \beta_{1S}m + \beta_2({}_S m - {}_L m) + \beta_{3L}$$
$$\qquad (20, 7)$$

$$\gamma_1 z + \gamma_2({}_S m - {}_L m) + \gamma_{3L}\dot m - \gamma_{4L}m = \delta_1 z - \delta_2({}_S m - {}_L m) -$$
$$- \delta_{3L}\dot m + \delta_{4\ L}m + \delta_5 \int_0 z\,dt \qquad (20, 8)$$

21. Let us investigate somewhat further the consequences of these equations, in order to find out whether they explain some of the most characteristic features in the cyclical behavior of interest rates. The equations are rather complicated: they are simultaneous equations; for ${}_L m$ even differential equations, with a right-hand member. In order to clarify the connections involved, we

may start with a very special case and gradually proceed to more complicated cases. Since statistical tests, to be discussed later (section 26), have suggested that the influence of the terms with $_L \dot m$ is weak, we start by omitting them. Further, we assume for a while that the other substitution terms (those with $_S m - _L m$) also vanish. This means considering the two markets as separate markets. The equations now become:

$$(a_1 + \beta_1)_S m = a_1 z.$$ (21, 1)

$$(\gamma_4 + \delta_4)_L m = (\gamma_1 - \delta_1) z - \delta_5 \int_0^t z \, dt.$$ (21, 2)

Two cases may be distinguished: (i) $\delta_5 = 0$ and (ii) $\delta_5 > 0$. In the former case, the two interest rates move parallel with each other and with z if $\gamma_1 > \delta_1$. Their amplitudes are proportional to

$$\frac{a_1}{a_1 + \beta_1} \cdot \text{and} \quad \frac{\gamma_1 - \delta_1}{\gamma_4 + \delta_4},$$ respectively. Since the numerator of the second

fraction is a difference, there is a good chance that the second fraction is smaller than the first and hence that $_L m$ shows less pronounced fluctuations than $_S m$. In non-mathematical terms, the explanation would be that, during the business cycle, both the demand and the supply schedules for long credits shift — in the same direction — whereas for short credits only the demand schedule shifts (the structure of the banks, determining the supply schedule, being constant) and hence the short-term rate shows larger fluctuations than the long-term rate. If, however, $\delta_1 > \gamma_1$, the fluctuations of the two rates would be opposite to each other, which does not fit the facts. Assuming (ii) that $\delta_5 > 0$ would mean the introduction, in the movements of $_L m$, of a component leading the general cycle by one-quarter of a period, and would therefore mean that m would show a more or less pronounced lead to that cycle, which does not fit the facts either.

Now let us no longer assume that the two markets are isolated, and study the consequences of the introduction of the substitution terms with $_S m - _L m$. We now get:

$$(a_1 + a_2 + \beta_1 + \beta_2)_S m - (a_2 + \beta_2)_L m = a_1 z.$$ (21, 3)

$$(\gamma_2 + \delta_2)_S m - (\gamma_2 + \gamma_4 + \delta_2 + \delta_4)_L m = (-\gamma_1 + \delta_1) z + \delta_5 \int^t z \, dt.$$

(21, 4)

These equations may be written in the simpler form:

$$_Sm - \theta_{1L}m = \theta_2 z \qquad (21, 5)$$

$$\theta_3{}_Sm - _Lm = \delta_1'z - \delta_5'\int_0^t z\,di \qquad (21, 6)$$

where all θ's are <1, since $\quad \theta_1 = \dfrac{\alpha_2 + \beta_2}{\alpha_1 + \alpha_2 + \beta_1 + \beta_2}$

$$\theta_2 = \dfrac{\alpha_1}{\alpha_1 + \alpha_2 + \beta_1 + \beta_2}$$

$$\theta_3 = \dfrac{\gamma_2 + \delta_2}{\gamma_2 + \gamma_4 + \delta_2 + \delta_4}$$

Moreover, $\delta_1' = \dfrac{\delta_1 - \gamma_1}{\gamma_2 + \gamma_4 + \delta_2 + \delta_4}$ and $\delta_5' = \dfrac{\delta_5}{\gamma_2 + \gamma_4 + \delta_2 + \delta_4}$.

Assuming again, for a moment, (i) δ_5' to be zero, the solution of the equations (21, 5) and (21, 6) may be written in the form:

$$_Sm(1 - \theta_1\theta_3) = (\theta_2 - \theta_1\delta_1')z$$

$$_Lm(1 - \theta_1\theta_3) = (\theta_3\theta_2 - \delta_1')z.$$

If now $\delta_1' < \theta_3\theta_2$, we have: $\quad \theta_3\theta_2 < \theta_2 \quad$ and $\quad \delta_1' > \theta_1\delta_1'$ from which it follows that $\theta_3\theta_2 - \delta_1' < \theta_2 - \theta_1\delta_1'$, or that the amplitude of $_Lm$ is less than that of $_Sm$, a result similar to that obtained formerly.

Assuming, then, (ii) δ_5' to be positive, we get an additional term in the solutions for both interest rates. It may, however, be proved that this term is relatively stronger for $_Lm$ than for $_Sm$; therefore we should still obtain a pattern not in conformity to reality; $_Lm$ would lead $_Sm$ and both would lead the general cycle. The same may be proved for the re-introduction of the $_L\dot{m}$-terms.

22. So far these results are none too good; some of the most characteristic features of reality, namely, that both the short and the long rate lag behind the general cycle and the long rate more than the short, have not yet been explained, even by the introduction of the dynamic terms with $_L\dot{m}$. There is, however, another dynamic element that has to be introduced; demand for long credits, as far as it takes the form of capital issues, ordinarily lags behind its determinants, since this type of demand is often provisionally satisfied by the use of anticipation credits: industrial demand by bank credits, government demand by the floating of treasury bills. Hence the terms $\gamma_1 z - \gamma_{4L}m$ in equation (20, 5)

may be lagged; e.g. over half a year. Taking this period as the unit of time, we write $\gamma_1 z_{-1} - \gamma_{4L} m_{-1}$ instead, and try to solve the equations proceeding, as before, from the simpler to some of the more complicated cases. Leaving out, to begin with, all substitution terms and the term with $\int_0 z\, dt$, we have:

$$(a_1+\beta_1)_S m = a_1 z \qquad (22, 1)$$

$$\delta_{4L} m + \gamma_{4L} m_{-1} = \gamma_1 z_{-1} - \delta_1 z. \qquad (22, 2)$$

The first of these equations again tells us that $_S m$ will fluctuate parallel with z; the second is more complicated, and we shall only solve it approximately. The left-hand member may be written as

$(\delta_4+\gamma_4)_L m_{-\theta}$,   where $\theta \sim \dfrac{\gamma_4}{\delta_4+\gamma_4}$;   the right-hand member as

$-\gamma_1(z-z_{-1})+(\gamma_1-\delta_1)\, z$. In this latter expression, the first term represents a component showing a lag of a quarter of a cycle approximately; the second either (if $\gamma_1 > \delta_1$) no lag or (if $\gamma_1 < \delta_1$) one of half a cycle. The result is that $_L m$ shows a lag behind the general cycle, which is either somewhere between $\frac{1}{4} T - \theta$ (where T is the length of the cycle) and $-\theta$ (i.e. a lead of θ) or somewhere between $\frac{1}{4} T - \theta$ and $\frac{1}{2} T - \theta$, both cases in the neighborhood of $\frac{1}{4} T - \theta$, since the coefficient $\gamma_1 - \delta_1$ of the z term is smaller than that of the $z - z_{-1}$ term. Thus we get a lag in, but a somewhat pronounced one.

Let us now introduce again the terms with $_S m - _L m$; this yields the equations:

$$(a_1 + a_2 + \beta_1 + \beta_2)_S m - (a_2 + \beta_2)_L m = a_1 z \qquad (22, 3)$$

$$(\gamma_2 + \delta_2)_S m - (\gamma_2 + \delta_2 + \delta_4)_L m = -\gamma_1 z_{-1} + \delta_1 z + \gamma_{4L} m_{-1} \qquad (22, 4)$$

As in section 21, we transform them into

$$_S m - \theta_1 \,_L m = \theta_2 z \qquad (22, 5)$$

$$\theta_3 \,_S m - _L m = -\gamma_1' z_{-1} + \delta_1' z + \gamma_4' \,_L m_{-1} \qquad (22, 6)$$

where θ_1 and θ_2 have the same meaning as before, $\theta_3 = \dfrac{\gamma_2 + \delta_2}{\gamma_2 + \delta_2 + \delta_4}$,

$$\gamma_1' = \frac{\gamma_1}{\gamma_2 + \gamma_4 + \delta_2} \quad \text{and} \quad \delta_1' = \frac{\delta_1}{\gamma_2 + \gamma_4 + \delta_2}.$$

The solution may now be approximated by the following procedure: provisionally we consider m_{-1} as given, and we deduce,

from (22, 5) and (22, 6):

$$_Sm(1-\theta_1\theta_3) = (\theta_2 - \theta_1\delta_1')z + \theta_1(\gamma_1'z_{-1} - \gamma_4'{}_Lm_{-1}) \qquad (22, 7)$$

$$_Lm(1-\theta_1\theta_3) = (\theta_3\theta_2 - \delta_1')z + (\gamma_1'z_{-1} - \gamma_4'{}_Lm_{-1}) \qquad (22, 8)$$

For (22, 8) we write:

$$(1-\theta_1\theta_3+\gamma_4')_Lm_{-\theta} = (\theta_3\theta_2 - \delta_1' + \gamma_1')z_{-\eta} \qquad (22, 9)$$

$$\text{where } \theta = \frac{\gamma_4'}{1-\theta_1\theta_3} \text{ and } \eta = \frac{\gamma_1'}{\theta_3\theta_2 - \delta_1'}.$$

The exact behavior of $_Lm$ can only be determined, on the basis of this approximate solution, when the various coefficients are known numerically. What we may show is, however, that now at least the possibility exists of a result in accordance with the real facts. If $\theta_3\theta_2 - \delta_1' + \gamma_1'$ is a small number, $_Lm$ will have a small amplitude, again chiefly because both the supply and the demand schedule shift upwards during the cycle. If further $\eta > \theta$, $_Lm$ is lagged behind z. The smallness of $_Lm$'s amplitude would make it legitimate to neglect, in (22, 7), the term with $_Lm_{-1}$; hence (22, 7) would yield us the solution for $_Sm$.

In addition, a comparison of (22, 7) and (22, 8) shows that $_Sm$ and $_Lm$ are both composed of two z terms, one without and one with a lag, of which the former is stronger for $_Sm$, whereas the latter is stronger for $_Lm$. This implies that both interest rates lag behind the general cycle, but $_Lm$ more than $_Sm$, in accordance with reality.[1]

23. We shall conclude our digest of recent literature on the formation of interest rates by a few paragraphs on the results of econometric analysis in this field obtained up to now. This short survey will be given the combined character of a review and of a program. What are the most conspicuous features of the level and the course of interest rates in the principal countries demanding explanation? They may be classified as follows:

1. For an interesting analysis of this relation of J. R. Hicks, Mr. Hawtrey on Bank Rate and the Long-Term Rate of Interest, The Manchester School of Economic and Social Studies X (1939), p. 27. P. Lorenz (Eine Differentialgleichung der Wirtschaftsforschung und ihr Integral, Blätter für Versicherungsmathematik und verwandte Gebiete, Beilage zur Zeitschrift für die gesammte Vers.-Wissenschaft, Bd. ?g. Heft 3, p. 212) has gone one step further and also proposed a "law" governing the process of mutual adaptation of the two rates, without, however, giving a clear-cut theoretical foundation of that law.

I. Differences in level.

Between countries. During the second half of the nineteenth century the levels of interest rates in the chief countries approached each other very closely; that of the United States having been considerably higher, at the middle of that century, than that of the European countries. After the first World War the rate of interest in Germany was considerably higher than elsewhere in western countries. Moreover, the rate of interest for first-rate bonds was in China very much higher than in a country like India, economically comparable to China; in India it does not surpass by very much the level in Great Britain.

Between various types of interest rates. As a rule, yield on first-rate bonds is lower than yield on stocks; there are, however, remarkable exceptions; for the United States after 1919 and Germany, 1870-1900.[2] Very large differences between rates for large credits to public authorites or concerns and for small credits to the population exist, notably in the far eastern countries.

II. Characteristics of movements.

Trend movement. The general trend throughout the nineteenth and twentieth century is a falling one.

Long cycles. Superimposed on this trend there are long cycles which coincide, roughly speaking, with those of the general price movement.

Shorter cyclical movements show themselves especially in short-term rates, where they are very pronounced. They lag somewhat behind the general cycle, as is expressed in the C-curve of the Harvard Barometer, and in addition they show the short "American" cycle (with a three- to four-year period) not only in the United States, but also in European countries. The long-term rate of interest practically does not show the normal cycle at all.

Influence of the World Wars. There is a pronounced rise in interest rates during and shortly after the first and no such rise during the second World War.

Seasonal and short irregular fluctuations may be left out of consideration in this essay.

24. As for the *differences in level*, they do not lend themselves so

2. Cf. Het verloop op langen termijn van de rentestanden, Statistische en Econometrische Onderzoekingen, ed. by Centraal Bureau v. d. Statistiek, The Hague, 1943, Nr. 2/3, p. 6.

well to econometric analysis as the characteristics of movements. Nevertheless, more could be done than has been accomplished so far. I do not know thorough studies on the high level of interest rates in the United States at the middle of the nineteenth century; the high level in Germany after the first World War, however, has found ample observation, albeit not along typically econometric lines of approach.[3] An econometric analysis of various sorts of interest rates in western and less developed countries has not been presented, as far as I am aware. Certainly it will be difficult to find adequate measures for the differences in risk premium that often will be necessary to explain the divergencies. Nevertheless, it does seem useful to embark upon a more systematic quantitative analysis of the phenomena enumerated.

25. As for the *characteristics of movements*, let us first consider somewhat more closely the *long-term movements* and apply our theoretical scheme to the explanation of these movements. This we may do by using a long period of reference; or to put it more exactly, to add up the market equation for a series of consecutive unit periods. This procedure will to some extent simplify matters. First, the "stock items" in the equation will largely cancel out against each other; taking as an example market equation (8, 1), $^2M_{t+1}$ for the first elementary period stands against $-^2M_{t+1}$ for the second elementary period, and hence cancels out. This need not be exactly so, since in the equation for the first period $^2M_{t+1}$ represents an expected value and in that for the second period an ex-post value. If we may assume that, *over a long period, there are no systematic differences between expectations and realizations, however,* we may neglect that fact. Hence, in the equation obtained by the adding up, for all elementary periods included in our long period of reference, of the market equation, we are only left with the stock quantities referring to the beginning and the end of the long period, and the equation obtained is of quite the same type as our elementary market equation. This implies, however, that the magnitude of the flow terms is *now much more preponderant* in comparison with that of the stock terms: the flow terms are approximately proportional to the length of the period of reference and the stock items are not.

A second consequence of the consideration of a long period

3. Cf. e.g. E. Welter, Die Ursachen des Kapitalmangels in Deutschland, Tübingen, 1931.

of reference is that items showing a rapid movement will lose in significance: the separate summands will, to some extent, cancel out against each other. Such may be the case, for example, with credit creation and with hoarding flows. These items will, during the course of one business cycle, alternatively show negative and positive values and their total will be rather small. The same applies to the elementary periods before and after a panic.

All this means an approach to "classical" or to static theories: the influence of the steady flows of savings, depreciation allowances, new investment, reinvestment purchases and sales of "old" securities and their determinants, becomes preponderant. But a complete approach to these theories is not possible; in pure statics, for example, cash holdings for risks do not make sense, but in reality they do, even in the long run.

An important feature of purely statical theory is that the *rate of interest is only dependent on "real" or physical, not on nominal, factors.* For example, a doubling of the money in circulation will, according to this type of theory, only lead to a doubling of nominal prices and money items, but not to a change in any "real" variables and hence not in the rate of interest. To put it another way, the rate of interest, being the ratio between the income of the capital owner and his capital, will not be influenced in the long run, since both his income and his capital will have doubled. It is therefore doubtful whether the long-term correlation between the price level and the interest rate can be easily understood. Macaulay[4] tried to do so by the following reasoning, which he unfortunately has not translated into a close quantitative approach. Starting from the hypothesis that the relevant credit contracts are long-term contracts, and the interest charges therefore remain constant if the general price level rises, he points to the increases in profits and hence in demand for further investments that follow. Against this increased demand he sees no adequately increased supply of credits, since incomes will not have risen so much. In view of the great increase in profits, this is not clear; it may be true, as Macaulay contends, that salaries and incomes of bond proprietors will not have risen so much, but incomes of entrepreneurs will.

Another attempt to account for the correlation between price

4. Some Theoretical Problems Suggested by the Movements of Interest Rates, etc., New York, 1938, p. 206.

level and interest rates, and at any rate theoretically sounder, although it may seem somewhat artificial, is the one undertaken by Irving Fisher.[5] Fisher points to the influence of a persisting *rate of increase* in the price level on the demand for and supply of credits. Such a rate of increase means a continuous depreciation of money in terms of goods, and hence of a capital that is lent. This will induce the lender to require a rate of interest that compensates him for this loss, whereas the borrower will be prepared to pay this higher rate: his productive activities will enable him to earn this additional charge. Hence it is to be expected that, once the rate of increase in prices has penetrated into the minds of the lenders and the borrowers, the rate of interest will be higher than normal during times of rising prices and lower during times of falling prices. If now, however, the correlation between interest rates and the price level itself is to be explained, it must be borne in mind that the rate of increase in prices precedes by one quarter of a period the price level itself. The correlation between interest rates and the rate of increase in prices *must be assumed to be one with a lag* and, in fact, of an average lag of about ten years. It would take ten years, then, before people would be aware of a long-run tendency of prices to rise or to fall. Is not this rather long? Five years may be accounted for immediately: in order to eliminate the ordinary cyclical movements we have to base our judgment on some moving average for about nine years, and it is only after five years that this can be known. Whether or not the remaining five-year lag can be understood psychologically is, to my mind, the point of doubt that arises here.[6] Further investigations on this point would certainly be interesting.

An influence of a similar character is the one, already alluded to, exerted by *the rate of increase in the interest rate itself*. If there is a persisting tendency in one or the other direction, this means that there is an opposite movement in the prices of bonds which, in the case of a fall, means a capital loss to the owner, and in the case of a rise, a capital gain. Correspondingly, the demand for bonds — i.e. the supply of credit — will be lower than it would have been otherwise in periods of rising interest rates, and higher in times of falling rates. A tendency to a rise in interest rates will therefore bring about a tendency to a high level. Since interest

5. The Theory of Interest, New York, 1930.
6. I presented this interpretation in the Dutch article referred to above.

rates, as we saw, move parallel to the general price level, this influence works parallel to the influence of a rising price level just discussed, and it will not be possible to separate these two forces.

Apart from these monetary long-run influences, what are the *physical determinants of the interest rate?* According to such theories as those of J. B. Clark and Von Böhm-Bawerk, and not in contradiction to our own scheme, it is the marginal productivity of capital: the factor behind the demand for credit. In the long run this marginal productivity will dpend on (*a*) the change in relative scarcity of capital in comparison with the other agents of production, chiefly labor; and (*b*) an eventual common change in the efficiency of all the agents. A remarkable econometric contribution in this field is that made by Douglas and his collaborators.[7] Douglas fitted a homogeneous first degree production function

$$P = cL^\lambda C^{1-\lambda} \qquad (25, 1)$$

to various sets of statistical material, where P is the quantity of industrial output, L the quantity of labor and C the quantity of capital engaged and c and λ are constants. He found λ to be about 2/3 or 3/4. From (25, 1) it is easily deduced that the marginal productivity of capital

$$\frac{\partial P}{\partial C} = (1 - \lambda) \left(\frac{L}{C}\right)^\lambda \qquad (25, 2)$$

where $\frac{L}{C}$ is the inverted relative scarcity of capital relative to labor. Since L and C are, on conditions to be enumerated, to be considered as data, it would follow that equation (25, 2) tells us what determines, in the very long run, the course of interest rates. They should move inversely with the 2/3d or 3/4th power of the relative scarcity of capital. Very roughly, the figures at our disposal seem to confirm this result.[8] The conditions to be fulfilled are that, in the long run, the supply elasticities of labor and capital are equal to zero; that, in other words, the remuneration of these agents does not change the proportion of the total supply available. Interesting investigations on these points have also been made by Douglas[9]; the last word on these questions has not yet been spoken.

7. P. Douglas, The Theory of Wages, New York, 1934, and various later publications.
8. Cf. my article quoted above.
9. Loc. cit.

26. Finally let us turn briefly to the problem of explaining the *cyclical movements of interest rates.* In sections 20–22 we have already given some attention to a possible theoretical set-up for such an explanation. The need for knowledge of the numerical values of the coefficients there introduced is evident: we demonstrated that the most conspicuous characteristics of cyclical movements may be explained only if these numerical values fulfill certain conditions. As far as I am aware, an investigation along exactly the lines indicated in sections 20–22 is not possible, because of lack of data on gross flows of new short credits. Statistical data now at hand only permit of a stock or a net flow analysis, as far as short credits are concerned.

An attempt to determine the functional relationship between the interest rates and the demand for and supply of credits along such lines has been made in my investigation on business cycles in the United States referred to above.[1] A few words may be added about the set-up chosen there in the light of the foregoing analysis. Three interest rates have been distinguished: two long-term rates, the yield on bonds m_{Lb} and the yield on shares m_{La} and one short term rate m_s. Accordingly, three demand and three supply functions have been determined. No more frequent figures than annual ones were available. For the supply and demand functions of short credits (being identical with demand and supply functions, respectively, for short "claims") a stock analysis was made; in the light of our discussion of this matter in section 19, we may hope this approaches a "gross flow" analysis, which we think is the appropriate way of analyzing this market. The correlations obtained are good, and do not therefore argue against this procedure. Of course, it would be interesting, however, to make further investigations in order to compare the results with those obtained by a gross and a net flow analysis. In addition, experiments with the use of smaller time units would be highly desirable.

The supply analysis of the long-term markets was also given the form of a stock analysis, i.e. the demand for bonds and shares was taken as a whole (demand for "new" as well as for "old" bonds and shares: the items $^1S_t' + {}^DE_t - {}^SE_t$ in the market equation (8, 1) were combined), and it was tacitly assumed that during the reference period — i.e. one year — all bonds and shares were scrutinized by their holders. Here the same remarks apply as

1. Cf. Chapter IV.

those made for the short-term market. The demand side of the long-term markets, however, was given the shape of a gross flow analysis: each year's issues were considered as a function of profits and interest rates then prevailing or some time before. This too seems legitimate, since the long period for which, from the side of the demand for these "credits" (or the supply of the corresponding "claims"), the contract is made makes it impossible each year to adapt their stock outstanding to the conditions of the market.

The procedure followed makes it difficult to trace the influence of the *savings process* on the rate of interest. It is assumed that family savings do not play any rôle in the supply of short credits. Hence they only enter into the picture at the supply side for long credits (or the demand side for long claims). Since here the stock analysis was followed, what had to come in was the annual figure of the *cumulation of anterior savings*. This is a series of figures showing mainly a trend and only a faint cyclical component. Its influence, therefore, cannot very well be separated from other influences showing mainly a trend. One further factor entering into this demand function for long claims and showing mainly a trend movement is the quantity of bonds and of shares present, at any moment, and since this also is a cumulation, the separation of the influences of these two factors by means of correlation analysis is practically impossible.

Among the rather scarce econometric investigations in the field of credit markets mention should also be made of the studies by A. J. Brown.[2] Here again a stock analysis is chosen.

<div style="text-align:right">J. TINBERGEN.</div>

NETHERLANDS SCHOOL OF ECONOMICS
  ROTTERDAM

2. The Liquidity Preference Schedules of the London Clearing Banks, Oxford Economic Papers 1 (1938), p. 49.

THE BRITISH IMPERIAL PREFERENCE SYSTEM

SUMMARY

Relation to the proposed International Trade Organization, 439. — I. Background of .the Ottawa Conference, 440. — The Ottawa agreements, 443. — Effects on Empire trade, 446. — II. Effects of Imperial Preference on the United States: United Kingdom, 451; Canada, 458; Australia, 461; New Zealand, 463; South Africa, 466; India, 467. — III. Conclusions, 468.

The suggested Charter for the proposed International Trade Organization made public by the State Department in September, 1946, and now under discussion at the trade and employment conference in Geneva would require members of the Organization to "enter into reciprocal and mutually advantageous negotiations . . . directed . . . to the elimination of import tariff preferences ' (Art. 18, par. 1). The Charter then spells out that "prior international commitments shall not be permitted to stand in the way of action with respect to tariff preferences," and that "all negotiated reductions on most-favored-nation import tariffs shall operate automatically to reduce or eliminate margins of preference." These provisions are identical in substantive content to the provisions of the Proposals for Expansion of Trade and Employment which were released in December, 1945, and of which the suggested Charter is "an elaboration."

While these provisions will affect the tariff preferences now in effect between the United States and Cuba and the Philippine Islands, and in the British, French and Dutch Empires, there is little question that they were directed primarily at the British Imperial preference system. The questions of continuance or discontinuance, and conditions for modification, of this system were among the principal items on the agenda of the Clayton-Keynes discussions held in late 1945. These were the discussions out of which came the now often repeated statement that "the Government of the United Kingdom is in full agreement on all important points in these proposals (Proposals for Expansion of World Trade and Employment) and accepts them as a basis for international discussion."[1] This "full agreement" covered, inter . alia, the preferential tariff provisions of the Proposals, which, as already

1. Joint Statement by the United States and the United Kingdom Regarding the Understanding Reached on Commercial Policy. December 6, 1945.

noted, were very similar to the preferèntial tariff provisions of the suggested Charter.

The heavy emphasis in this country on modification, and if possible elimination, of the British Imperial preference system is a reflection of the views generally expressed by United States government spokesmen, members of Congress and the foreign trade fraternity that this system was responsible for serious curtailment of United States exports to the British Empire during the 1930's, and that unless it is now modified, it will prove a difficult stumbling block to any possible future expansion of United States exports to the Empire.

At the same time, many segments of British public opinion are equally convinced that the British people would lose more than they would gain by modification of the preferential system. They maintain that the system was of great benefit to British foreign trade during the 1930's, that it helped to expand intra-Empire trade, and that any modifications would endanger the United Kingdom's current drive to expand exports to 50–75 per cent above prewar volumes.

In their references to prewar experiences the contentions on both sides of the Atlantic are generally more notable for their emotional and political content than for their reasoning based upon economic and statistical evidence. This paper, therefore, reviews briefly the nature of the Ottawa agreements which established the Imperial preference system, and then examines the results of the operations of the system from two specific points of view: (1) what were the effects of the system on British foreign trade during the period 1932–1938; (2) to what extent did the system hinder United States exports to the British Empire during this same period?

The period 1932–38 is chosen for analysis because this was the only period during which the effects of the system were relatively clear-cut. After 1938 the trade of the Empire was conditioned first by preparations for war and then by the war itself, by the operations of the sterling area, and by Lend-Lease. These factors minimized the importance of the preference system during the war years.

I. THE OTTAWA AGREEMENTS

Background. The drive for both general tariff protection to British producers and tariff preferences for the Empire has a

historical development tracing back to the 1890's. In fact, in 1897 Canada took the first practical step in the direction of preferences, when she granted special rates to Empire suppliers. Between 1903 and 1907 New Zealand, South Africa and Australia established preferential rates ranging from five to fifteen per cent for Empire products.

Before World War I, a number of the Dominions requested the United Kingdom to establish some form of preferential tariff treatment for their products. In order to grant such preferences, Britain would have had to permit Empire goods to enter duty free and to tax the imports of non-Empire goods, or to levy a small tariff on Empire goods and a higher one on goods from non-Empire sources. In those days, however, Britain was essentially a free-trade nation, and these requests were refused.

During World War I, the United Kingdom established the McKenna duties on a relatively small number of items. Although originally established primarily for revenue raising purposes, these duties were subsequently looked upon by the affected industries as "protection," and they strenuously objected to all attempts to reduce or eliminate them. The motor car industry was one of those affected.

The Finance Act of 1919 established preferences for Empire products on commodities then subject to import duties (McKenna duties). Clause 8 stated that "with a view to conferring a preference in the case of Empire products, the duties of the goods specified shall . . . be charged at reduced rates (hereinafter referred to as "preferential rates") . . . where the goods are shown . . . to have been consigned from and grown, produced or manufactured in the British Empire." Subsequently, the Key Industries Duties, 1921, provided for import duties for a period of five years on a small group of commodities when imported from non-Empire sources, but left imports from Empire sources on the free list. These duties were renewed by the Finance Act of 1926.

As a result of these developments, preferences were established for Empire suppliers of automobiles, alcoholic spirits, some industrial goods, sugar and tobacco. For example, foreign automobiles were subject to a duty of $33\frac{1}{2}$ per cent, while Empire produced automobiles were subject to a duty of only $22\frac{2}{3}$ per cent. Actually, however, this particular preference was not really a preference, since none of the Dominions or other parts of the

Empire then produced automobiles for export. In fact, one of the results of this "preference" was the establishment in Canada of subsidiary plants of United States automobile producers.

In the overall, however, these preferences did not have any significant effects on the sources of British imports. As of November, 1931, the great bulk of imports entering the United Kingdom were still duty free.

During 1930 and 1931 United Kingdom and Dominion exports suffered heavily, as did the exports of all other major exporting nations. The Dominions were affected because they were essentially producers of raw materials, the prices of which were sharply reduced in the world's markets. British exports, on the other hand, were primarily of manufactured goods, e.g. textiles, machinery, industrial products and other fabricated goods. The prices for these goods did not decline as much as the prices of raw materials, but foreign demand was reduced as a result of reduced purchasing powers and reduced demands.

In 1931, after the formation of the "National" Government under Ramsey MacDonald, the United Kingdom adopted full-fledged tariff legislation for the first time. The Abnormal Importations (Customs Duties) Act, 1931, and the Horticultural Products (Emergency Customs Duties) Act, 1931, empowered the President of the Board of Trade and the Minister of Agriculture and Fisheries, respectively, to impose duties up to 100 per cent ad valorem on a wide range of foreign manufactured goods and horticultural products, provided the Chancellor of the Exchequer concurred and the House of Commons subsequently approved. Empire products were specifically exempted. Duties of 50 per cent were subsequently imposed on a small number of commodities, including clothing, various textiles and radio equipment

The Import Duties Act, 1932, enacted in the early months of that year provided for direct duties of 10 per cent ad valorem on a broad range of commodities when imported from non-Empire sources. The rates could be increased on the recommendation of a Tariff Advisory Commission which the Act created. The products of the Dominions, India, Southern Rhodesia and territories under mandate to the Dominions were specifically exempt from duties until November 15, 1932. After that date, their tariff treatment would depend on the results of the Imperial Economic Conference, which had already been scheduled for July 21, 1932.

Products of other parts of the Empire were exempt without time limitations. Shortly after the enactment of this Act, the 10 per cent rate was increased for many commodities to from 15 to 33 per cent.

During this same period, the various parts of the Empire were pressing the United Kingdom to grant them preferences. The legislation noted above gave the United Kingdom a powerful bargaining weapon. If the Dominions and other territories would grant special preferences to her exports of manufactured goods, she could grant them preferences in her markets for their raw materials. At Ottawa, Britain's principal efforts were directed, therefore, to increasing the markets for her exports. The Dominions, however, were all engaged in granting protection to their own manufacturing industries. The chief competitor for their own manufactures was the United Kingdom, except in the case of Canada, whose principal competitor and supplier was the United States. Therefore, if Britain were to grant Empire preferences, it was likely that she would retain the existing duties on foreign products and continue to permit Empire products to enter duty free, or impose preferential duties for Empire products, or grant quantitative preferences to Empire suppliers The Dominions, on the other hand, were more likely to increase their existing duties on foreign goods and to give British exports the benefits of the spread between the existing ones and the newly increased ones than to reduce the existing duties on British goods.

The Agreements. At Ottawa the United Kingdom made five principal concessions.[2] (All Empire imports were subsequently given comparable preferences)

1. All Empire products which, at the time of the conclusion of the agreements, were admitted duty free would continue to be admitted duty free. This covered about 80 per cent of all imports from the Empire.[3] The rest would be subject to revenue duties or to pre-1931 protective duties.

2. The United Kingdom agreed to impose new or revised duties on a specified list of goods coming from foreign countries, e.g. butter, various fruits, eggs, wheat, linseed, unwrought copper, etc.

2. Imperial Economic Conference, Agreements Concluded at Imperial Economic Conference, 1932, Command paper.
3. Benham, Frederic, Great Britain Under Protection, p. 92.

3. The United Kingdom agreed to maintain the general 10 per cent ad valorem duties imposed by the Import Duties Act and not to reduce these duties except with the consent of the Dominions, India, etc.

4. The United Kingdom agreed to restrict the quantities of chilled and frozen beef, frozen mutton and lamb, and bacon and ham, to be imported from foreign countries. However, it reserved the right to restrict imports of similar goods from Empire sources if these entered in too large quantities. This reservation was designed to protect domestic producers.

5. The United Kingdom agreed to maintain specific margins of preference on a small number of articles from each of the Dominions, India, etc., and not to reduce these margins except with the consent of the affected Empire area. For example, Empire tobacco producers were guaranteed special preferences for ten years, and the coffee preference was increased.

The agreement with Canada stipulated that if any country were to attempt by state action to frustrate the intention of preferences, imports from that country of the affected commodities were to be prohibited until such time as the margins of preference were reëstablished. Canada insisted on this clause, in order to prevent her timber exports to the United Kingdom from being undersold by Russia.

The Dominions and India agreed to the following:[4]

1. *Canada.* At the time of the Ottawa conference, Canada was applying a general surcharge of three per cent on all imports, and was converting £ sterling into Canadian dollars at a higher rate than that ruling in the open market.[5] Canada agreed to abolish the surcharge on British imports as soon as the finances of Canada would allow and to bring the official exchange rate into line with open-market rates. It was agreed that tariff protection would be given only to industries which were "reasonably assured of sound opportunities for success" — a very loose promise — and to keep tariffs sufficiently low to enable efficient British exporters to compete. A number of British goods were to be admitted duty free. Where the British already enjoyed preferential treatment, the margins of preference were not to be reduced. On about two

---

4. Imperial Economic Conference, op. cit., and Canada, Bureau of Statistics, Empire Tariff Preferences, 1932.

5. Benham, Frederic, *Great Britain Under Protection*, p. 92.

hundred additional items preferences for the United Kingdom were increased by either reducing the tariff on British goods or raising the tariff on foreign goods, or both.

2. *Australia.* Before Ottawa, Australia had one of the highest tariff schedules in the world. In addition, in 1930, the Commonwealth prohibited the import of some goods (generally luxury goods), imposed a surcharge of 50 per cent on the import of some other goods, and levied a general tax, known as a primage duty, on all imports. In July, 1931, this general tax was 10 per cent. At Ottawa, Australia agreed to abolish the import prohibitions and surcharges on British goods and to remove the primage duty as soon as financial conditions made it possible. (This duty was never completely removed. However, the Australians did provide for a differential in favor of the British in applying this duty.) Australia also agreed to protect only those industries which appeared reasonably sure of survival under competition. Preferences on British goods were granted, the margins ranging from 15 to 20 per cent. (In some cases the margins eventually ran as high as 65 per cent.) This was generally done by increasing the tariff on foreign goods. Subsequently, Australia established a three-column tariff schedule — preferential, intermediate and general. United States exports were subject to the general rates, i.e. the highest rates.

3. *South Africa.* Prior to Ottawa, South Africa had a rather moderate tariff schedule. Preferences on some commodities were granted to the United Kingdom and to other British territory which granted reciprocal treatment to South Africa, generally about five per cent. At Ottawa, South Africa agreed to continue these preferences and to grant new ones on a short list of manufactured items, mostly in the iron and steel and machinery classes. This was done mainly by increasing the tariff on foreign goods. South Africa eventually established a three-column tariff schedule — minimum, for most British goods; intermediate; and maximum. Imports from the United States were subject to the maximum rates.

4. *New Zealand.* Before Ottawa, New Zealand had a moderate tariff schedule, but granted considerable ·preference to British goods. In addition, in 1931, the Dominion imposed a new and higher schedule, but still retained the British preferences, and imposed new surtaxes and primage duties. At Ottawa, New Zea-

land agreed to abolish the surtax on British goods, to waive the primage duty on goods otherwise duty free as soon as financial circumstances made it possible to do so, and to maintain existing preferences. She made a commitment similar to that agreed to by Canada and Australia to protect only such industries as showed reasonable prospects of survival under competition.

5. *India.* Prior to Ottawa, India had a tariff schedule under which rates on some goods were as high as 50 per cent. British goods enjoyed very few preferences. At Ottawa, India agreed to grant the United Kingdom preferences to 10 per cent on a large number of manufactured goods, and 7½ per cent on automobiles, omnibuses, and accumulators. This agreement could be terminated on six months' notice.

In the years following the Ottawa conference, the Dominions and India signed reciprocal agreements with each other which extended the preferential system to their exports to each other. Similarly, other British territories also were granted, and in turn granted, preferences.

Effects of the Agreements. Tables I, II and III summarize the shifts in Empire trade that resulted from Ottawa.[6] From Table I it appears that United Kingdom exports to Empire countries increased from 43.5 per cent of total British exports in 1930 to 49.9 per cent in 1938, a difference of only 6.4 per cent. If this fact is considered in relation to the value of British exports, the percentage of shift assumes greater importance. The years 1931–38 were years during which all exporting nations suffered, years during which all the major exporters had more capacity to produce for export than was actually exported. Thus, if it is assumed that the United Kingdom had the physical capacity to export more in value to non-Empire countries, but did not because imports were down, the same logic would lead to the conclusion that if the United Kingdom had not increased the proportion of exports to Empire countries, British exports would have decreased even more than they did. By increasing the proportion of her exports to the Empire, the United Kingdom, therefore, benefited by the preferential system, which was undoubtedly responsible, in large

6. These figures, from the Statistical Abstract for the British Empire, 1938, vary slightly from percentage figures given in the Annual Yearbooks of the Dominions. Quoted in Benham, Frederic, Great Britain Under Protection, pp. 256, 257, 262.

measure, for these changing proportions in the direction of United Kingdom exports.

TABLE I

PROPORTIONS OF UNITED KINGDOM EXPORTS TO EMPIRE COUNTRIES

	1927	1930	1931	1932	1934	1935	1937	1938
Canada........	4.1	5.1	5.3	4.5	5.0	5.0	5.3	4.8
Australia........	8.6	5.6	3.7	5.5	6.6	6.9	7.2	8.1
New Zealand....	2.8	3.1	2.9	2.8	2.9	3.1	3.9	4.1
South Africa.....	4.3	4.6	5.6	5.0	7.6	7.9	7.9	8.4
India...........	12.0	9.3	8.3	9.3	9.3	8.9	7.5	7.7
Irish Free State..	5.1	6.0	7.8	7.1	4.9	4.7	4.1	4.3
Rest of Empire .	9.2	9.8	10.1	11.1	10.6	11.5	12.4	12.5
Total Empire....	46.1	43.5	43.7	45.3	46.9	48.0	48.3	49.9
Total Non-Empire	53.9	56.5	56.3	54.7	53.1	52.0	51.7	50.1

TABLE II

PROPORTIONS OF UNITED KINGDOM IMPORTS FROM EMPIRE COUNTRIES

	1927	1930	1931	1932	1934	1935	1937	1938
Canada.........	4.5	3.6	3.8	6.1	6.9	7.4	8.7	8.6
Australia........	4.3	4.4	5.3	6.6	6.8	7.2	7.0	7.8
New Zealand....	3.8	4.3	4.4	5.3	5.5	5.0	4.9	5.1
South Africa	1.7	1.9	1.5	2.2	1.6	1.8	1.7	1.6
India..........	5.4	4.9	4.3	4.6	5.8	5.4	6.3	6.1
Irish Free State..	3.5	4.1	4.2	3.8	2.4	2.5	2.1	2.5
Rest of Empire...	6.7	5.9	5.2	6.8	8.1	8.3	8.7	8.7
Total Empire....	30.1	29.1	28.7	35.4	37.1	37.6	39.4	40.4
Total non-Empire	69.9	70.9	71.3	64.6	62.9	62.4	60.6	59.6

It would be a mistake, however, to overemphasize the importance of these increased proportionate exports to the Empire. In the first place, the principal increases were registered for Australia, New Zealand, and South Africa, while decreased proportions are recorded for India, which had formerly been the United Kingdom's largest Empire customer, and the Irish Free State. Exports to Canada remained relatively stable in percentage terms. In terms of commodities, the proportion of British exports

TABLE III

TRADE OF THE EMPIRE WITH GREAT BRITAIN

A. Percentage of Total Imports Coming from the United Kingdom

	1930	1931	1932	1933	1934	1935	1936	1937	1938
Canada	16.1	17.4	20.7	24.4	22.1	21.2	19.4	18.2	17.6
Australia	39.6	40.6	42.5	43.4	43.6	41.5	43.6	42.1	42.5
New Zealand	47.4	49.0	49.8	51.3	50.4	50.3	49.4	49.6	47.9
South Africa	46.9	45.5	46.3	50.3	48.8	48.7	46.3	42.4	43.6
Irish Free State	80.0	80.8	76.6	69.9	66.7	72.4	53.2	50.0	50.5
India	39.2	36.7	37.6	41.8	41.1	39.3	39.0	31.5	31.4

B. Percentage of Total Exports Going to the United Kingdom

	1930	1931	1932	1933	1934	1935	1936	1937	1938
Canada	26.7	28.5	36.4	39.9	41.5	42.1	42.1	40.2	40.6
Australia	44.9	50.3	47.6	47.7	52.8	49.9	50.0	46.1	51.2
New Zealand	82.1	89.7	90.2	87.3	82.3	84.5	80.6	76.0	83.8
South Africa	47.9	43.4	42.1	37.9	41.1	42.5	40.9	41.1	40.7
Irish Free State	92.1	96.3	96.1	93.9	93.4	91.7	91.5	90.7	92.6
India	23.4	27.4	27.8	32.0	31.6	31.4	32.4	32 4	33.7

to Empire countries increased for iron and steel products, electrical machinery and apparatus, other machinery, motor cars and parts, cotton and woolen piece goods, and chemical drugs, but decreased for agricultural implements and machinery, and other metals and manufactures. Table IV sets forth the percentages of British exports of specified commodity groups to Empire countries.

The most striking feature, however, of the results of the Ottawa preferences, was the sharp increase in the proportion of United Kingdom imports from the Empire (Table II). From 1930 to 1938 these increased from 29.1 per cent to 40.4 per cent of total United Kingdom imports. The principal beneficiaries were Australia, Canada, and the smaller parts of the Empire. In this case, there is little question that the Dominions and most of the rest of the Empire benefited, especially so since the intervening years were years of declining raw material prices and overall world surpluses of most raw materials and foodstuffs.

TABLE IV

PROPORTION OF BRITISH EXPORTS OF COMMODITY GROUPS
TO EMPIRE COUNTRIES

	1930 %	1938 %
Iron and Steel Products...........................	50.8	62.8
Agricultural Implements and Machinery............	62.5	54.5
Electrical Machinery and Apparatus...............	56.6	71.7
Other Machinery	43.1	54.7
Motor Cars and Parts............................	62.7	69.7
Other Metals and Manufactures...................	37.5	35.5
Yarn and Thread, including Wool Tops.............	22.5	26.4
Cotton Piece Goods..............................	49.4	55.8
Woolen Piece Goods.............................	28.1	34.0
Apparel, except Footwear........................	68.8	67.9
Other Textiles..................................	55.3	59.9
Chemicals, Drugs, Dyes, etc......................	50.0	58.0

Source: Statistical Abstract for the British Empire, 1938.

For example, in 1930 United Kingdom imports of bacon and ham were 10.2 million cwts., of which 50 per cent came from Denmark, 14 per cent from the United States and about two per cent from Canada. In 1938, 7.5 million cwts. were imported, of which 45.3 per cent came from Denmark, 5.3 per cent from the United States, and 20 per cent from Canada. British imports of chilled and frozen beef in 1930 were 11.6 million cwts., of which 72.4 per cent came from Argentina, 9.5 per cent from Uruguay, 6.9 per cent from Australia, and 2.6 per cent from New Zealand. In 1938, 12.3 million cwts. were imported (the cost was less than the 1930 imports, which were greater in quantity) of which 60 per cent came from Argentina, 5.7 per cent from Uruguay, 18 per cent from Australia, and 7.3 per cent from New Zealand.

There were important shifts in the sources of supply for three additional principal United Kingdom imports of foodstuffs: frozen mutton, butter and wheat. In 1930, frozen mutton imports totalled 6.4 cwts., of which 23.4 per cent came from Argentina, 6.2 per cent from Uruguay, 52.5 per cent from New Zealand, and 12 per cent from Australia. 1938 imports were 6.9 million cwts., of which Argentina supplied 12.9 per cent, Uruguay 2.9 per cent, New Zealand 53.6 per cent and Australia 27.5 per cent.[7]

With respect to the imports of meat, it is important to note

7. Statistical Abstract for the British Empire, 1938.

that the diversion from foreign sources to Empire sources was due only secondarily to tariff preferences. The prime causes of the changes in the sources of supply were the quantitative limitations that were imposed on imports from non-Empire sources, and the licensing system that was put into effect. These latter techniques were, of course, other forms of preference.

Imports of butter in 1930 were 6.8 million cwts. Denmark supplied 32.3 per cent, New Zealand 23.5 per cent, and Australia 14.7 per cent. In 1938, when imports reached 9.5 million cwts., Denmark supplied 25.2 per cent, New Zealand 27.4 per cent and Australia 18.9 per cent.[8]

The shift in supply sources in the case of wheat is shown in Table V.

TABLE V

UNITED KINGDOM IMPORTS OF WHEAT

Total Imports	104.8 million cwts.	101.6 million cwts.
	1930	1938
Australia..........	25.0%	30.5%
Canada...........	12.1	38.4
British India......	3.1	4.3
United States ...	20.0	15.5
Argentina.... .	14.5	5.6 *

Source Statistical Abstract for the British Empire, 1938

British imports of non-foodstuffs did not show such notable shifts as imports of foodstuffs. This was due primarily to two factors: (1) the Empire countries were not, by and large, large exporters of manufactured goods, so that British imports of such goods did not indicate any great shifts from non-Empire to Empire countries; (2) British imports of raw materials other than foodstuffs before Ottawa generally came from Empire sources, e.g. tin, rubber, copper, manganese, jute, sisal, petroleum. This trend continued after Ottawa.

There are two major effects of these shifts in the sources of supply of British imports which are frequently overlooked. First, the decreased United Kingdom imports from non-Empire sources reduced the amount of sterling available to these sources, and thus reduced their ability to purchase British goods. In this sense, British exports cannot be said to have benefited from the Empire

8. Ibid.

preference structure; they were simply re-directed, not increased. In fact, it may be said that this re-direction tended to decrease British exports because some of the additional sterling thus made available to the Empire sometimes was used to amortize debts and to repurchase British-held investments in Empire, whereas if this same amount of sterling had been available to non-Empire countries, it would, it is likely, have been used to purchase additional British goods.

Second, if Table I on the proportions of United Kingdom exports to Empire countries is examined in connection with Table III, part A, on percentages of total imports of Empire countries coming from the United Kingdom, it will be noted that, although the United Kingdom increased the proportions of her exports to the Empire, the principal Empire importers did not increase the proportions of their imports from the United Kingdom. In fact, South Africa, India and the Irish Free State decreased the percentages of their imports coming from the United Kingdom and only Australia registered any appreciable increase. This indicated two things: first, that the greatest benefits of the Empire preference structure were gained by the Dominions in the form of increased proportionate exports to the United Kingdom at a time when the world's raw materials and food markets were very low; second, that the United Kingdom was becoming more dependent on the Empire for markets, while at the same time the dependence of the Empire on the United Kingdom as a source of supply was actually diminishing. In other words, the direction of British exports was becoming less flexible, while the Dominions and India were acquiring greater flexibility in choosing their sources of supply. The latter was the exact opposite of what the British had intended to accomplish at Ottawa.

II

EFFECTS OF IMPERIAL PREFERENCES ON THE UNITED STATES

There are a number of factors which should be considered in any evaluation of the effects of the British Imperial preference system on Anglo-American trade, or on United States trade with other parts of the Empire. This is due to the fact that the shifts in United Kingdom sources of supply cannot be said to have been caused by the preferential structure alone. In the first place,

slight, and in some cases, major shifts of sources of supply have occurred in the past without the intervention of preferences. For example, technological changes have had very vital effects on the direction of trade. Secondly, the worldwide economic depression in the 'thirties had sharp effects on the distribution of United States exports even to those areas where tariff preferences were absent. Thirdly, in the period that followed Ottawa, British lending to, and investments in, non-Empire countries virtually ceased, whereas lending to and investments in the Empire continued. Fourthly, exchange rate fluctuations during the 'thirties, including the devaluation of the pound sterling in September, 1931, and the subsequent devaluation of the dollar in 1934, generally worked to the advantage of the British — in this context — and thus provided British exporters with a competitive advantage over their American confrères.

Another fact which tended to limit the effects of the Ottawa agreements on the United States was that United States exports to the British market did not — with certain exceptions — compete with British imports from Empire sources. These exceptions were in foodstuffs, tobacco, and timber. The various parts of the Empire were not, generally speaking, exporters of manufactured products which competed with United States exports; in other cases they exported commodities which did not compete with the United States, e.g. rubber, tin, tea, cocoa. However, there is some evidence to indicate that the Dominions took advantage of preferences to try to build up industries. This was particularly true of Canada, which obtained much of the capital for these developments from the United States.

Finally, it must be noted that concepts of curtailment and impairment of trade vary. Some United States exporters argue that the extent of curtailment of United States exports to Empire markets is best measured by actual trade statistics. Others maintain that the Imperial preference system, by placing United States exports at a competitive disadvantage, curtailed *potential* exports, and thereby did severe damage to United States foreign trade. There is some logic in this last contention. However, since the curtailment of these *potential* exports cannot be measured with any exactness or even probability, it is necessary for purposes of evaluation to analyze the actual trade statistics.

United Kingdom. Available statistics indicate that in the

overall the proportion of United States exports to the United Kingdom after Ottawa did not vary very sharply from pre-Ottawa exports. Table VI gives total United States exports from 1926–38 and the proportions of these exports imported by the United Kingdom. These figures lead to the conclusion that the United States exported roughly the same proportion of total exports to the United Kingdom after Ottawa as before.

TABLE VI

TOTAL UNITED STATES EXPORTS AND PROPORTION TO THE UNITED KINGDOM
1926–1938

Year	Total Exports ($ million)	Per Cent to the United Kingdom
1926–1930	4,773.3	17.5
1931–1935...........	2,025.2	18.4
1936................	2,456.0	17.8
1937................	3,349.2	16.0
1938.......	3,094.4	16.8

Source: Statistical Abstract of the United States, 1944.

TABLE VII

PROPORTION OF UNITED KINGDOM IMPORTS FROM EMPIRE COUNTRIES,
NON-EMPIRE COUNTRIES, AND UNITED STATES

Year	Per Cent Empire	Per Cent Non-Empire	Per Cent United States
1927..	30.1	69.9	17.3
1930...........	29.1	70.9	15.0
1931...........	28.7	71.3	12.2
1932...........	35.4	64.6	12.0
1934...........	37.1	62.9	11.2
1935...........	37.6	62.4	11.5
1937...........	39.4	60.6	11.1
1938...........	40.4	59.6	14.0

Source: Statistical Abstract for the British Empire, 1938.

The United Kingdom, however, decreased the percentage of its imports coming from the United States in the years following Ottawa. This is shown in Table VII. It will be noted that the decline in the percentage of United Kingdom imports from the United States began even before Ottawa. Further, despite the continued decline in the total proportions of United Kingdom

imports supplied by non-Empire countries, the United States maintained its ratio of United Kingdom imports, and even increased it in 1938.

Ham.   It has already been indicated that the principal competition from Empire countries for United States exports in the United Kingdom market affected foodstuffs, tobacco, and timber.   In the case of meats, the diversions in sources of supply were accomplished not so much by the new preferences as by the quantitative limitations that were imposed on imports from non-Empire countries, including the United States.   Ham was the principally affected United States meat export.   In this case, therefore, the United States bent its efforts to getting the British to agree to modify the percentage of total imports from non-Empire sources that could be supplied by the United States.   This involved breaking down the old British classifications, which grouped bacon and ham together, and getting a larger quota established for United States exports of ham.   This was accomplished in 1938.

Wheat.   Wheat and fruits were the other major foodstuffs affected by the Ottawa agreements.   In the case of

TABLE VIII

UNITED KINGDOM WHEAT IMPORTS AND PERCENTAGES
FROM EMPIRE COUNTRIES AND UNITED STATES

Year	Total British Imports (000 bushels)	Per Cent from United States	Per Cent from Empire
1927............	206,147	32.3	47.1
1928............	193,344	22.8	51.0
1929............	208,632	19.9	36.0
1930............	195,580	20.1	40.5
1931............	222,915	9.4	42.6
1932............	197,189	4.4	67.2
1933............	209,767	*	67.0
1934............	191,567	0.1	56.1
1935............	188,955	0.6	54.1
1936............	188,108	*	83.7
1937............	180,803	3.6	65.4
1938............	189,703	15.6	63.2

* Less than 0.05 per cent.
Sources: Annual Statements of the Trade of the United Kingdom; Accounts Relating to Trade and Navigation of the United Kingdom, August, 1939. Quoted in Kreider, C. J., The Anglo-American Trade Agreement, p. 93.

wheat, United States exports to the United Kingdom shrank to almost nothing from 1933 through 1936 (Table VIII). It should be noted that there were severe droughts in the United States in 1934 and 1936, and that the wheat crop of 1935 was smaller than usual. This reduced the ability of the United States to export any substantial quantities. Further, the principal Empire competitor in the United Kingdom market was Canada. This was, and is, because Canada is the only Empire producer and exporter of hard wheat, the same type exported by the United States. Subsequently, under the terms of the Anglo-American Trade Agreement, 1938, the United Kingdom preference for Empire wheat was abolished.

Fruits. The principal concessions made by the United Kingdom at Ottawa on fruits were the establishment or modifications of preferences on fresh apples and pears, canned fruit salad, pineapples and grapefruits. Of these, the most important from the United States point of view was fresh apples, of which the United Kingdom imports from the United States in 1929 were valued at almost $20 million. In the five years preceding Ottawa, the United States had supplied from 36.2 per cent to 62.1 per cent of total British imports, with Empire countries supplying almost all of the balance. As in the case of wheat, the principal Empire competitor of the United States was Canada. Australia and New Zealand were the other Empire suppliers, but the growing and shipping seasons in these Dominions are different from the growing and shipping seasons in the United States and Canada. After Ottawa, the share of United Kingdom imports from the United States declined to 18.3 per cent in 1933, rose to 31.8 per cent in 1935, and then declined again to 22.1 per cent in 1938. The Anglo-American Trade Agreement reduced the preference by reducing the specific tariff rates applicable to imports from the United States.

With respect to citrus fruits, United States exports to the United Kingdom of oranges and grapefruits fell sharply after Ottawa. In the case of oranges, however, the principal United States competitors were not Empire sources, but rather Spain and Brazil. In the case of grapefruits, the United States' share of total British imports fell from 87 per cent in 1927 to 23 per cent in 1936. Part of this decline was probably due to

the preference granted South Africa and the British West Indies, but the most important reason was the tremendous growth of Palestine as an exporter to the United Kingdom. Palestine, it may be noted, did not receive any preferences.

Tobacco. In the five years prior to Ottawa, the United States supplied from 78.9 per cent to 85.6 per cent of total United Kingdom imports of unmanufactured tobacco. At Ottawa the United Kingdom agreed not to reduce the margin of Empire preferences for ten years, except with the agreement of the Empire countries affected. The tariff rates applied to imports from both non-Empire and Empire sources were very high, ranging from 9s. 6d. to 10s. 6½d. per pound from non-Empire sources, with Empire sources getting a preferential margin of about two shillings per pound. (From 1934 to 1938 the pound sterling-dollar rate fluctuated from $4.889 to $5.039.) However, even after Ottawa the British consumer continued to prefer United States tobacco (Table IX).

TABLE IX

UNITED KINGDOM TOBACCO IMPORTS AND PERCENTAGES
FROM EMPIRE COUNTRIES AND UNITED STATES

Year	Total British Imports (1,000 pounds)	Per Cent from United States	Per Cent from Empire
1927..	222,265	79.8	18.4
1928......... .	217,785	78.9	19.8
1929..........	240,026	85.6	13.1
1930.............	237,028	83.4	14.9
1931........ ...	194,046	81.0	17.7
1932............	174,912	71.6	27.3
1933.......... ..	211,108	75.7	23.3
1934	238,875	79.1	19.9
1935	251,620	80.5	17.9
1936	270,899	78.8	19.4
1937.	267,460	76.0	21.4
1938	344,858	74.4	24.0

Sources Annual Statement of Trade and United Kingdom, 1931, 1934, 1937; Accounts Relating to Trade and Navigation of United Kingdom, August, 1939. Quoted in Kreider, C J , The Anglo-American Trade Agreement, p. 101.

Although the United States endeavored to get the British to reduce the preferential margin, the British could not do so because of the commitment made at Ottawa not to lower the

margin of preference for ten years. Southern Rhodesia, which was the principal Empire country affected, refused to agree to a lowering of the preferential margin.

Timber. At Ottawa, the British agreed to permit Canadian exports of timber to the United Kingdom to enter duty free and to establish a 10 per cent duty on foreign timbers "of all kinds imported into the United Kingdom," when such timbers were also imported in substantial quantities from Canada. The rate on foreign timbers could not be reduced, except with the consent of the Canadian Government. Since most of the timber imported into the United Kingdom from Canada was Douglas fir, this provision directly affected United States exports to the United Kingdom, because the principal United States export was also Douglas fir.

After Ottawa, United States exports of timber to the United Kingdom declined sharply, while Canadian exports increased very substantially (Table X). Finland, which was the major supplier before Ottawa, continued to be the major supplier. While there is some question as to whether the preferential margin was wholly responsible for the decline in United States exports and the increase in Canadian exports, there is little doubt that it was a major factor in this development.

TABLE X

UNITED KINGDOM IMPORTS OF SAWN AND PLANED SOFTWOODS

(1,000 standards)

Year	Total United Kingdom Imports	Imports from United States	Imports from Canada	Imports from Finland
1929.....	1,915	112	80	493
1930.....	1,790	128	83	389
1931.....	1,602	80	55	375
1932.....	1,538	43	77	385
1933.....	1,997	48	207	557
1934.....	2,338	43	410	636
1935.....	2,124	44	330	630
1936.....	2,521	46	449	661
1937.....	2,570	57	491	584
1938	1,985	35	457	470

Source. Compiled from data gathered by William Brandt's Sons & Co , Timber Department, "Approximate Statistics Exports and Imports of Sawn and Planed Softwoods" 1933 and 1938   Quoted in Kreider, C J , The Anglo-American Trade Agreement, p. 96.

During the discussions on the Anglo-American Trade Agreement, United States negotiators tried to get the United Kingdom to agree to lower the margin on Douglas fir. After much discussion, the British agreed to rearrange their tariff classifications, but they made no clear-cut concessions.

Canada. Canada first introduced preferences for imports from the United Kingdom in 1897. This preference was extended from time to time to other parts of the Empire and to reciprocating Dominions. After Ottawa the system became generally applicable, by agreements, with almost all parts of the Empire.

In 1935, three years after Ottawa, Canada signed a Trade Agreement with the United States which had the effect of substantially narrowing the preferential margins established at Ottawa. Then, in 1938, a new and much more comprehensive agreement was signed. Under this agreement the United States granted Canada concessions on 202 items and sub-items covering 83 per cent of Canadian exports to the United States in 1937. On 107 of these items, the United States made the maximum reduction of 50 per cent allowed under the Trade Agreements Act. Of the remaining items, 58 were accorded reductions ranging from 10 to 50 per cent, five were assured of continuance of the existing duties, and 32 were assured of continuance of entry free of duty. The principal Canadian products affected were lumber, shingles, horses, cattle, dairy products, hog products, potatoes, fish, certain grains, poultry, pulp and paper, metals, non-metallic minerals, and some manufactured goods. In return, Canada made concessions to the United States on 447 tariff items and sub-items, covering 57 per cent of total Canadian imports from the United States in 1937. Reductions were made on 283 items, and assurances were given that existing rates on 146 additional items would not be increased. Canada also agreed to remove the special excise tax of three per cent then being levied on all of these items. The Agreement contained safeguarding clauses as to quantitative restrictions, customs valuation, variations in rates of exchange, and preventing the "principal benefit of a concession going to a third country." Both countries agreed to extend unconditional most-favored-nation treatment to each other, but with reservations for Canadian Empire preferences and the United States preferences

to Cuba, the Philippine Islands and the Panama Canal Zone.[9]

As a result of these agreements, the preferential margins agreed to by Canada at Ottawa were sharply reduced. Thus, at the outbreak of the war in 1939, about 40 per cent of total Canadian imports consisted of articles free of duty from *all* countries. About one per cent consisted of articles which were dutiable from both British and non-British countries, but without preference. Approximately 60 per cent of all imports were subject to preferential treatment of one kind or another.

Of the 60 per cent of total imports which were subject to preferential treatment, approximately two-fifths of those from British countries consisted of commodities which were free of duty when imported from British countries but dutiable when imported from other sources. Including those commodities which came in duty-free from British countries, but not including commodities which were free of duty from all countries, the preferential rates averaged about two-fifths of the non-preferential rates. In 1939, when the average rate for United States goods was about 23 per cent, the average rate for British goods was about 9.5 per cent.

Immediately after Ottawa, the share of the United States in total Canadian imports decreased from the 1920–29 levels, while the share of the United Kingdom and other British countries increased. After the signing of the Trade Agreements, however, the share of the United States tended to increase, that of the United Kingdom to decrease. The share of other Empire countries increased in both cases. This is evident in Table XI.

An important consideration in analyzing the effects of the Ottawa preferences on United States exports to Canada is the composition of United States and United Kingdom exports to Canada. In the first place, many imports from the United States enter Canada duty-free. These are raw materials and foodstuffs which the United Kingdom does not export. Second, Canadian consumers continued to look to the United States for supplies because of the proximity of the supply source, the similarity of requirements, and the fact that many Canadian producers who imported parts and semi-finished items were subsidiaries of United States companies. Third, the period immediately after Ottawa was a period of great flexibility of exchange rates and prices, due to the devaluation of the pound sterling in September, 1931, and

9. The Canada Yearbook, 1940, pp. 492–493.

TABLE XI

SHARE OF UNITED STATES, UNITED KINGDOM AND OTHER EMPIRE COUNTRIES
IN CANADIAN IMPORTS

Year	Total Canadian Imports (\$ million Canadian)	Per Cent from United States	Per Cent from United Kingdom	Per Cent from other Empire Countries	Per Cent from all Empire Countries
1922	747.8	69.0	15.7	4.3	20.0
1926	927.3	65.6	17.6	4.9	22.5
1929	1,265.7	68.6	15.3	5.0	20.3
1930	1,248.3	67.9	15.2	5.1	20.3
1933	406.4	57.2	21.3	8.3	29.6
1934	433.8	54.9	24.2	8.2	32.4
1937	671.9	58.6	19.3	10.2	29.5
1938	799.1	61.0	18.2	11.0	29.2

Source. The Canada Year Book, 1940, pp 513, 515, 526

the subsequent devaluation of the United States dollar in 1934. United States prices, even after devaluation, were in many instances much higher than comparable Empire quoted prices. This probably induced some shifts from United States to Empire sources. fourth, after Ottawa, many United States producers circumvented the intent of the preference system by building plats in Canada. This had the dual effect of (1) decreasing United States exports to Canada, and (2) giving the Canadian subsidiaries the benefit of preferences in other Empire markets. This, in turn, tended to increase Canada's share of Empire imports and to decrease the share of the United States.

As a rough average, it may be said that Canadian imports from the United States after Ottawa were divided as follows: 30 per cent raw materials (including unprocessed food), six per cent partly manufactured commodities, and about 64 per cent fully manufactured goods. The comparable percentages for the United Kingdom were about 8.5 per cent, 10 per cent, and 81.5 per cent. The percentages for other Empire countries were 35.0 per cent, 37 per cent and 28 per cent, respectively. However, the Empire countries' (other than the United Kingdom) exports of raw materials and semi-finished products to Canada were generally of products which did not compete with United States exports, e.g. rubber, tin, jute, cocoa, etc. It is particularly interesting to note that before 1930 Canada used to import tin and rubber via the United

States, and that the increased proportion of imports from Empire countries after Ottawa was partially due to the fact that after 1930 these products enter Canada directly from the producing areas. This was an additional reason for the declining share of the United States in Canadian imports immediately after Ottawa.

It is likely, however, that preferences did play a rôle in the declining share of the United States in Canadian imports of raisins, some cotton textiles, and chemical products. In the overall, however, other reasons, as already noted, contributed more to this declining share than did preferences, particularly after the Trade Agreements were signed in 1935 and 1938. In fact, it may be said that the principal effects of Imperial preferences on United States exports, insofar as they concerned Canada, came from the benefits Canada gained as against the United States in other Empire countries, particularly in the United Kingdom. These advantages have already been described in the section on the United Kingdom.

Australia. In 1932 Britain and Australia extended preferential treatment for the other's produce: Britain guaranteed to Australia for three years (later extended) free entry of a number of specified commodities, and preferential treatment on others; Australia granted to the United Kingdom preferential treatment on a wide range of commodities, provided these commodities would not compete with domestically produced merchandise. It was further provided that the benefits of the British Preferential Tariff could be extended wholly or in part to any British non-governing colony, protectorate or mandate — but not to the Dominions. Imports from all other areas were made subject to a general tariff which ranged from 50 to 75 per cent more than the tariff imposed on British merchandise.

In 1936 this system was revised, in order to include a third tariff schedule — the Intermediate Tariff. The purpose of this, said the Minister for Trade and Customs, was to provide "a convenient avenue for expressing the level of duties which the Government propose should form the basis of trade treaties. The rates proposed under the protective items of the Intermediate Tariff express, in every case, a protective level for Australian industry, as well as preserving the margins required under the Ottawa Agreement."[1] Three gradations of tariff schedules were thus provided for. Significantly, the merchandise of the United

1. Official Yearbook of the Commonwealth of Australia, 1940, p. 239.

States and Japan — except for cotton piece goods and artificial silk — were placed on the highest tariff schedule.

In addition, in 1936, Australia instituted a "Trade Diversion Plan." This plan had as its aims:[2]

1. An improvement in the Australian defense positions as a result of further development of secondary industry, particularly iron, steel and engineering, and complete motor car construction in Australia. (This last aim was never achieved.)

2. An improvement in the Australian financial position, since an improved trade balance was expected to facilitate the payment of Australia's overseas interest and dividend bill.

3. To improve the Australian bargaining position in the United Kingdom.

4. To increase exports of primary products.

It was intended that imports would be obtained from those countries which were good customers of Australia and which, it was hoped, would become even better customers as Australia increased her purchases from them. The Commonwealth Government proposed to accomplish this program in two ways:[3]

1. Adoption of a special licensing system over a carefully selected range of imports.

2. Imposition of higher duties where this course appeared more desirable.

All British goods except motor car chassis were exempt from the license restrictions. In actual operation, however, licenses were granted freely to those countries, whether Australia had a favorable or unfavorable balance with them, where the conditions of trade were satisfactory to Australia.

The implementation of this plan resulted in a series of license restrictions which, because of the commodities affected, seemed particularly directed against United States exports to Australia. This country retaliated — under the provisions of the Trade Agreements Act, 1934 — and withdrew the most-favored-nation treatment previously accorded to the Commonwealth.

In the face of the restrictive measures, United States imports from Australia, excluding gold, fell from £stg.8,727,000 in 1936–37 to £stg.2,695,000 in 1937–38, from 7.36 per cent to 2.39 per cent

2. *Ibid.,* p. 762, Harris, H. L , Australia's National Interests and National Policy, p. 100.

3. Official Yearbook of the Commonwealth of Australia, 1940, p. 762.

of Australian exports. However, United States exports to Australia *increased* from £stg.12,959,149 to £stg.17,758,684, from 14.64 per cent to 16.26 per cent of total Australian imports. This was primarily due to the nature of the United States exports.

During the 1930's the principal United States exports to Australia were motor-power machinery, machine tools, metal working machinery and other producer goods, motor car chassis and parts (as a development of assembly plants in Australia), chemicals and drugs, and certain raw materials in which Australia had a complete or near complete deficiency, and in which the United States had a dominant export position — cotton, sulphur, aluminum, lubricating oil, tobacco. Very few manufactured or consumer items were normally imported from the United States. Of these, the principal items were paper and stationery, films, some apparel and textiles, office machinery, and optical, surgical and scientific instruments. Britain generally supplied the bulk of Australia's import requirements for these consumer goods. British-American competition in Australia's import markets was, during this decade, principally confined to drugs and chemicals, machinery, and motor vehicle chassis and parts.

In the capital goods field, Great Britain supplied about 52 per cent (1938–39) of total Australian imports of metals, metals manufactures and machinery, as compared with the 19.8 per cent supplied by the United States. In drugs and chemicals, the United States supplied almost 18 per cent, as compared with 40.8 per cent from the United Kingdom. There is reason to believe that, had Australia not granted Empire preferences and attempted the trade diversion plan, United States exports of these products would have constituted a larger share of Australia's imports.

In one other product — softwood timbers — there is some evidence to indicate that United States exporters lost part of their market in Australia. In this case, Canada was the beneficiary.

Subsequently, in the summer of 1943, the Australian Government decided to apply the Intermediate Tariff rates on United States merchandise.

New Zealand. New Zealand first introduced Imperial preferences in 1903. At first only a few items of Empire origin were benefited, but in 1907, 1921, 1927, 1930 and 1934 the coverage was widened considerably. At Ottawa, New Zealand agreed to preserve the existing margin of preference on United Kingdom goods

when the margin of preference did not exceed 20 per cent, and where the margin exceeded 20 per cent not to reduce it below 20 per cent without the consent of the United Kingdom Government.

The range of preference may be better understood by analysis of the rates applied to all imports in 1938. In that year 23 per cent of total imports were admitted duty-free from all sources, while eight per cent were dutiable at equal rates from both British and non-British sources. This left 69 per cent of total imports subject to preference.[4]

It should also be noted that in 1938 New Zealand instituted both import and exchange control. The reasons were to "ensure that after overseas debt commitments have been met from the sterling funds, the maximum funds will be provided for the importation of essential commodities, with particular regard to the needs of primary and industrial production in the Dominion."[5]

TABLE XII

PROPORTIONS OF NEW ZEALAND'S IMPORT SOURCES OF SUPPLY

	1930	1932	1934	1936	1937	1938	1939
United Kingdom........	47.6	51.0	50.4	49.4	49.6	47.9	46.8
Australia..............	6.8	9.6	10.3	11.2	11.7	12.9	12.9
Canada...............	8.9	4.5	6.7	7.5	8.1 ˙	8.8	8.8
Total British Countries.	68.4	71.2	73.8	72.7	73.5	73.6	74.2
Japan.............	1.3	1.9	2.7	3.0	2.9	2.2	2.1
Netherlands East Indies..	1.9	3.5	4.4	4.0	4.0	3.9	4.9
United States..	17.8	14.8	12.0	12.7	12.4	12.4	11.4
Total Foreign Countries	31.6	28.8	26.2	27.3	26.5	26 4	25 8

Source: New Zealand Official Yearbook, 1941, p 231.

As may be seen from Table XII, some interesting changes occurred in the sources of supply for New Zealand's imports. (1) The United Kingdom's share of the New Zealand market increased slightly immediately after Ottawa, but then declined steadily until the outbreak of the war. (2) The Canadian share of the New Zealand market remained relatively stable. (3) Almost all of the increase in the share of the New Zealand market gained by Empire

4. New Zealand Official Yearbook, 1941, p. 213.
5. Ibid., p. 222.

countries went to Australia. (4) Most of the decline in the share of the non-Empire countries was suffered by the United States. (5) Despite the declining share of all non-Empire countries, the Netherlands East Indies and Japan increased their share of the New Zealand market.

The decline of the United States share was due primarily to the decline in New Zealand's imports of American motor vehicles and petroleum products. With respect to motor vehicles, the United Kingdom first benefited, but in the latter half of the 'thirties Canada gained most of the former United States share. Canadian exports were, it may be noted, from branch plants and subsidiaries of the major United States producers. The petroleum products trade was lost to the Netherlands East Indies. In addition to preference, other factors contributing to the overall decline of the United States share of the New Zealand market were according to official New Zealand statements, "the high rate of exchange against New Zealand, particularly before the United States currency measures of 1933 and 1934, and in 1939 the policy of import control."[6]

The sharp rise in the Australian share of the New Zealand market is officially explained by the New Zealand Government as having been caused by a number of factors completely unconnected with the preference system. Significantly, almost all of this gain was at the expense of the United Kingdom. The growth of the iron and steel industry in Australia enabled the Commonwealth "to compete effectively in the New Zealand market with the manufactured products of the older countries."[7] The comparative nearness of Australia to New Zealand is a very strong factor operating to the benefit of Australian firms in the New Zealand import market, as are the closeness of personal contact possible between importers and the Australian manufacturers, the short period required for delivery of goods, and the intimate knowledge of New Zealand requirements, all acting strongly to the mutual advantage of the Australian exporter and the New Zealand buyer. Moreover, the New Zealand market presents a very strong attraction to Australian firms, since the potential demand in that country is of sufficient size to represent a very profitable addition to the home market. This stimulus does not act with anything like com-

6. Ibid., p. 231.
7. Ibid., p. 233.

parable force in the case of United Kingdom and continental manufacturers.

South Africa. Prior to Ottawa, South Africa extended preferential treatment to the United Kingdom and to any Dominion, colony or possession that granted reciprocal treatment to South Africa. Although the tariff schedule was in three columns, minimum (preferential), intermediate (applicable to most-favored-nations) and maximum (to all other countries), the preferential spread was not very great.

At Ottawa, South Africa signed agreements giving preferences to the United Kingdom, Canada and Eire. Although no specific agreement was signed with New Zealand, certain New Zealand produce was given the benefit of preferential rates. None of these rates, however, provided for a substantial preferential margin.

In addition, South Africa had specific anti-dumping duties covering not only f.o.b. costs, but also freight and exchange dumping. The last covered depreciated currencies. These anti-dumping duties applied to specific commodities and to specific sources of supply. In 1940 the United States, United Kingdom and Canada were among the countries to which anti-dumping duties were applied for specific commodities. Such duties were applied against the United States for rubber hose, wheelbarrows, glucose, piping and tubing, hauling and hoisting chains, and pneumatic tubes and tires. In the case of the United Kingdom, anti-dumping duties were applied on cement, rubber hose, piping and tubing, rope, coal-burning stoves and pneumatic tubes and tires. Canada was covered for cement, rubber hose, piping and tubing, hauling and hoisting chains and pneumatic tubes and tires.

As shown in Table XIII, the Empire as a whole and the United Kingdom in particular supplied a declining share of total South African imports during the 'thirties. On the other hand, non-Empire countries, and in particular the United States, provided an increasing share of South African imports. These figures lead to the conclusions (1) that the preferences granted to United Kingdom and the other Empire countries had no effect at all in shifting South Africa's sources of supply, the exact opposite of what the United Kingdom had intended to accomplish at Ottawa having occurred; and (2) that the preference structure had no ill effects on United States exports to South Africa.

TABLE XIII

PROPORTIONS OF UNITED KINGDOM, EMPIRE COUNTRIES
AND UNITED STATES IN SOUTH AFRICAN IMPORTS

	1930–34 Yearly Average	1935–39 Yearly Average
Total Imports................. ...	£SA 52,658,362	£SA 90,312,565
Per Cent United Kingdom..........	47.6	42.9
Per Cent All Empire Countries......	58.0	54.3
Per Cent United States.......	14.4	18.5
Per Cent All Non-Empire Countries..	42.0	45.7

Source Official Yearbook of the Union of South Africa, 1941, p 921

India. The effects of the granting by India of preferential treatment to the United Kingdom and to other Empire countries were analyzed by the United States Department of Commerce in 1939 in the following terms:[8]

> Provisions of the Ottawa Agreement affected about 55 per cent of India's normal imports from the United States, although in most instances it cannot be said that the effects have been seriously damaging. For example: in automobile imports, British preferences were raised from 38 to 45 per cent, with — in comparison with other factors — but little effect on sales of American automobiles; on machinery and millwork from the United States, preferences accorded Empire countries have had little or no effect, exchange rates being the only difficulties encountered in that field; on iron and steel products, corrugated culverts and some other lines were not affected, but other galvanized sheets felt the application of the increased duty; on tubes, pipes and fittings, however, the 10 per cent preferences given to Empire products offered serious competition to American business in these lines.

Although the Ottawa Agreements appeared satisfactory to Indian business men immediately after their signing, much adverse criticism was heard in the late 'thirties. Indian exporters maintained that exports to Empire countries did not expand as had been expected, and that preferences accorded Empire products in the Indian market built up prejudice in non-Empire countries against Indian products.

It is significant that despite preferential treatment, the United Kingdom's share in Indian imports declined from 45 per cent in

8. United States Department of Commerce, The United States in India's Trade (Trade Promotion Series No. 200), p. 68.

1928–29 to 38 per cent in 1936–37. (In 1937 Burma was separated from India and trade between these areas was henceforth recorded as exports and imports. This throws the 1937–38 and subsequent import percentages out of line.) The United States provided seven per cent of Indian imports in 1928–29 and six per cent in 1936–37. Significantly, the greatest gains in the Indian market were made by two non-Empire countries, Japan and Germany. In 1928–29 these two countries supplied seven and six per cent, respectively, of total Indian imports; in 1936–37, the corresponding figures were 17 and 10 per cent.

III. CONCLUSIONS

1. The system of Empire preferences built up by the British at Ottawa did not have the results desired by the British. While these preferences contributed to increasing the share of United Kingdom exports to Empire markets, they did not increase to any substantial extent the proportions of Dominion and Indian imports from the United Kingdom. In fact, the United Kingdom share of the Indian and South African markets decreased sharply. The trend in New Zealand also showed a progressive decline, while in Australia and Canada the shares remained relatively stable or increased slightly. Thus, while the United Kingdom was becoming more dependent on the Empire as a market, the principal Empire markets were becoming less dependent on the United Kingdom as a source of supply.

2. The greatest benefits of the preference system were gained by Canada and Australia in the form of their increased share of the United Kingdom market. However, other Empire exporters of raw materials and foodstuffs were also benefited, in that they maintained or increased their exports to the United Kingdom at a time of declining world demand and prices for the products they had available for export.

3. The United States was not greatly injured by the Imperial preference system. Only a few products were affected: bacon, tobacco, timber, and fruits to the United Kingdom; prunes, some textiles and chemicals to Canada; timber to Australia; motor cars to New Zealand. Despite the preference structure, the United States maintained its overall share of the United Kingdom and Indian markets, increased its share in the South African and Australian markets. It lost ground in the New Zealand market,

the smallest of the major Empire markets. The total share of the United States in the Canadian market also decreased slightly, but this was probably due more to other factors than to preferences.

4. The fact that the United Kingdom increased the share of imports from the Empire may actually have had a depressing effect on British exports to non-Empire markets. This shift in trade put more sterling in the hands of Empire countries, some of which was used to repay debts and to repurchase British-held Dominion and Indian securities and investments. Had a similar amount of sterling been made available to non-Empire countries, it is quite likely that it would have been spent for purchases from the United Kingdom.

5. The preference system could aid the United Kingdom and other parts of the Empire — if at all — only during a period of declining world trade and depressed raw material and foodstuff prices. In a period of rising world trade, such as is now hoped for, the preference system would actually be a hindrance to greater participation of the United Kingdom and other Empire countries. It would prevent them from getting the full benefits of increasing technological efficiency of non-Empire countries and of a freer multilateral system of trade. It is to the United Kingdom's and the Empire's advantage, therefore, to seek an expansion of world trade, a lowering of general tariff levels, and a removal of discriminatory trade practices, including preferences.

6. The United Kingdom's current international financial position is such that any attempt to increase Empire trade solely on a basis of preferences might be self-defeating. It might prevent an expansion of British trade with non-Empire countries at a time when Britain is urgently in need of non-Empire currencies and means of financing reconstruction.

7. There are, however, many additional factors to be taken into account in any evaluation of the United Kingdom's current trade and international financial positions. These include her actual capacity to produce for export, the availability of dollar markets for her own and the exports of other British countries, the recent price rises in the United States, which have had the effect of decreasing the physical volume of goods which the British may purchase with the $3,750,000,000 loan from the United States, the fear of depression in the United States, the agreement to convert freely all sterling acquired on current account after July 15, 1947,

and the disposition of the sterling balances which were accumulated by the members of the sterling area during the war. Each of these factors is likely to have just as significant influence upon the composition and direction of British trade as the preference system.

8. Even if it is assumed that the preference system is maintained and that world trade does not expand, it is questionable whether the United Kingdom would derive any considerable benefits from the preference system. Canada, Australia, South Africa and India have expanded their productive capacities considerably. In a period of curtailed world trade, it may be reasonably expected that these countries will endeavor to protect their domestic industries against both Empire and non-Empire competitors. In fact, all of these countries have already indicated that they may take this course of action, particularly if substantial results do not emerge from the current trade and employment conference in Geneva. Under such circumstances, they could still maintain preferences for the British but could so increase the tariff on both British and non-British goods as to limit the sale of both in their markets.

9 In the light of the increasing industrialization of the Dominions and India and of the United Kingdom's current financial position, and if it is assumed that substantial results do not emerge from the current conference in Geneva, and further, that the world is once more plunged into depression, the United States could be seriously affected by the preferential system and by the many other forms of trade barriers The rest of the world might then attempt to insulate itself from the possible effect of the export of competitive goods and unemployment from the United States.

DAVID L. GLICKMAN.

WASHINGTON, D. C.

INDUSTRIAL DEVELOPMENT OF THE NEAR EAST

SUMMARY

Beginnings, 471. — Turkey, 473. — Egypt, 476. — Iraq, 479. — Mandated territories: Syria and Lebanon, 480, Palestine, 484. — Developments during the war: Turkey, 488, Egypt, 490; Palestine, 491; Iraq, 492; Syria and Lebanon, 493. — Postwar problems: favorable factors, 494; the question of prices, 495; productivity, 495; frozen "war balances," 496; other problems, 498. — The present outlook, 499.

BEGINNINGS

Industry in the Near East has developed only recently. Such production as existed previously was carried on in small workshops by artisans and craftsmen, whose working personnel consisted primarily of the members of their own families. Under foreign influence some attempts had been made in Old Turkey to introduce industry, but without much success. Similar efforts met with no more success in Egypt, where both the rulers and foreign capitalists tried to build up factories. Mohammed Ali in the first half of the century and the Khedive Ismail later built industrial establishments and encouraged the flow of foreign capital toward the same end; but these factories for sugar, spinning, armament, woolen goods, etc. were for the most part abandoned, since the manufactures did not find a market or a sufficiently developed economy for their activities. The same fate was to a certain degree shared by the industries developed during World War I, when sea lanes were blocked and the demand for commodities grew. The armistice and the return to normalcy resulted in a sharp crisis in all this production and wiped out a great part of it.[1]

The trend throughout the world in most of the interwar period was toward economic self-sufficiency, tariff barriers, and building up home industries, and this tendency was strengthened in the crisis of the 1930's. The Near Eastern countries were exporting raw materials and importing finished goods. With the development of the world economic crisis the difference in prices on the world market between raw materials and industrial products widened, to the disadvantage of those countries exporting raw materials. This

[1]. Ahmed S. el-Amary, "La Structure economique de l' Egypte," L'Egypte Contemporaine, Vol. XVIII, No. 166–69 (1937), p. 196.

put a financial strain on the resources of such countries and resulted in a tendency to expand their domestic industrial production. In the Near East these tendencies were further augmented by non-economic reasons, namely, the growing nationalistic trends and the desire of the newly created states, or those striving for independence, to become economically independent.

Industry in the Near East was introduced at a time when other countries — Europe, the United States, Japan — already had developed industries, which were supplying foreign countries with cheap commodities. Moreover, this period is significant for the practice of dumping by industrial countries, when commodities were sold on foreign markets, not only below the market price of the country of production, but even below their cost of production.

The advantages of a cheap labor supply were more than offset by the necessity of training it and by the high overhead costs connected with the building of new industries. There was hardly a middle class in these countries which could build industries as a continuation of its activities as artisans or traders. Unlike the European countries during the Industrial Revolution, the artisans' workshop in the Near East did not furnish the needed labor force for the factories. In addition, all these countries lacked the necessary capital for industrial development. Industrial production therefore could not develop unless the Government supported it by financing, by an appropriate credit policy, and by protective tariffs.

The countries which were under mandates of the League of Nations — Iraq up to 1932, Syria and Palestine — were, however, at a disadvantage,[2] both because no local government was there to foster industrialization and because of the international situation in which these countries found themselves. The "open door" clause inserted in the mandates (Article 11 of the Treaty between Great Britain and Iraq 1922, Article 11 of the Mandate for Syria, Article 18 for Palestine) to prevent these countries from making bilateral agreements exposed them to the influx of foreign commodities. The mandatory countries (Great Britain and France), moreover, were more interested in keeping the Near Eastern

---

2. Turkey was also handicapped for the first few years, since the Treaty of Lausanne bound her to low duties on imports In 1929, however, Turkey repudiated her old treaties based on these limitations and launched a new policy.

countries as sources for needed raw materials (cotton, oil, silk) and markets for their goods than in helping these countries to produce manufactured goods themselves.

TURKEY

Turkey's intention of building up industry was motivated by strategic political reasons as well as by economic ones. The aim was to concentrate the population and create a greater number of dense areas which would not be vulnerable to any attack, at the same time spreading the industrial production over a number of centers in order to make it harder for the enemy to paralyze production. Another objective was to raise the standard of living of the population, both the agricultural, by creating a market for their produce, and the other sectors, by creating employment in industry and commerce. But the more vigorous promotion of industry was also motivated by the unfavorable balance of trade. This reached its highest level in 1929.[3] In the following years, with the disruption of the world market and the decline in prices of raw materials, which were the main exports of Turkey, the country was inclined to turn more and more to the production of consumer goods from her own raw materials. Clearing agreements were made with foreign countries whereby Turkey bought from each country only what she could pay for with her exports. At the same time she tried to produce goods in the country itself.

In the years 1923-33 new railway lines were built (2,600 km) and a number of national banks founded; the Ish Bank, founded in Ankara in 1924, was viewed as a means of drawing in local savings for investment purposes. Local raw materials were also cultivated. In 1926 a special institute for cotton was founded, whose task was to improve the kinds of cotton and supervise the development of its growth. Through this supervision and help, Turkey, which before World War I grew only a small quantity of cotton of inferior quality, began to increase her crops. In the same year the cultivation of sugar beets was started, and the Government attempted to encourage the development by offering tax reductions and other inducements to private growers.

Industrialization was also fostered at the beginning by en-

3. The excess of imports over exports amounted to £T. 101 million in 1929, as against 50 million in the preceding year. Since 1930 Turkey has had a favorable balance of trade.

couraging private inititative through grants, reduction of taxes and transport charges, granting of premiums and obliging the authorities to use the local products. This was stipulated in the Law for the Assistance of Industry, passed in 1927. Since the premiums and reductions were abused for private gains, the law underwent some changes, and parts of it, mainly the provisions for premiums, were abandoned entirely.

The number of enterprises which enjoyed state support rose from 342 in 1923 to 1,473 in 1932. The number of workers in all state-supported establishments was 69,000 in 1934. Almost half of these enterprises belonged to the category of agriculture, while about 300 were textile enterprises employing 16,000 people. The measures taken by the government were not, however, sufficient to bring about large-scale developments.

Turkey's industrialization by state means really began in 1934, when the first Five Year Plan was inaugurated. This provided for state planning and encouragement, for rigid controls of imports, and granted privileges to industries, including allotment of state lands, exemption from taxation, and reductions in freight. Credits were secured from Soviet Russia for the purchase of machinery, and Russian technicians were brought in to train a Turkish skilled labor force. In the further stages of industrialization credits were also secured from Great Britain and Germany for the purchase of equipment and machinery.

As the main method of promoting industrialization the state used the banks. The Sumer Industrial Bank, the Eti Mining Bank and the Ish Bank, endowed with state means, have been the main factors in the execution of the Five Year Plan inaugurated in 1934. They have been directly responsible for the foundation of new factories, for their organization and management.

The Sumer Bank was founded in 1933 for the purpose of establishing and managing state factories, granting credits to private industries, assisting in the training of qualified technicians, carrying on research and investigating industrial development. It received an annual grant from the budget of the state. By 1939 the Bank had developed the following enterprises: the Textile Works in Bakirkoy (1,173 workers), Kaysserie (4,500), Mazilli (2,400), Eregli (1,500), Merino Wool Mills (1,000), Hereke Wool Mills and Bunyan Wool Yarn Mills. In addition it owns the Izmir Paper and Cardboard Mills, Beykoz Shoe Factory, Tosya Rice

Cleaning Plant. It is also connected (together with other banks) with factories for the production of iron, steel, concrete, etc.

The Eti Bank was established in 1935 for the purpose of developing the mining and steel industry, its capital being contributed by the state in yearly installments from the Budget. It financed and managed the great steel and iron works at Karabuk, the Ergani Copper Mines, and the Heracle Copper Mines, as well as a number of electric power projects.

The Ish Bank is a semipublic institution founded in 1924 on the initiative of Ataturk. It derives its capital largely from private savings, but the funds are being invested in industrial enterprises. It participated in the construction of textile mills at Malatya, and in a bottle and glass factory.

Turkish Sugar Refineries have been established by the Sumer, Ish, and Agricultural Banks. The Turkish Tobacco Company has been set up as a state monopoly with the participation of the Ish and Agricultural Banks.

Before the beginning of the war a new Five Year Plan was started, which included the building of a large port in Zonguldak, the coal center, and fifteen new large factories (cotton, sugar, preserves, electrical power stations, synthetic gasoline, etc.) with an investment of eighty million £T.[4]

An idea of the rise in production of the Turkish industry may be gathered from the figures in Table I. The greatest development

TABLE I

OUTPUT (IN TONS) OF TURKISH INDUSTRIES[5]

	1927	1935	1939
Cotton goods........	2,500	12,300	22,800
Woolens............	564	2,355	....
Sugar..............	5,180	55,000	94,000
Cement.............	41,200	176,569	300,000

occurred in the textile industry, which employed 189,000 spindles in 1938, as against a few thousand in 1931.

We may also gauge the growth of all industrial production from the rise in the number of people employed in industry, which rose from 299,000 in 1927 to 656,000 in 1935.[6] The distribution of

4. One £T amounted to approximately $.80
5. Commercial Directory of Turkey, 1937.
6. Statistical Handbook of Middle Eastern Countries (Jerusalem, 1944), p. 131.

these workers in the various branches of production was as follows (in 1937):

	%
Mining................................	2.8
Building and Construction................	14.8
Wood.................................	10.0
Metal.............................. ..	11.0
Textile................................	12.8
Clothing..............................	15.7
Food.................................	13.2
Power and Electricity....................	0.2
Various and Unspecified.................	19.5

These figures should be regarded as indicative only, since over 100,000 workers are listed as "unspecified industrial workers." Subject to these limitations, we see that building and construction was the largest branch, followed by clothing, food, and textiles, while metal production employed somewhat more than one-tenth of the workers.

EGYPT

Attempts to build up an industry in Egypt go back, as we have seen, to the nineteenth century. Mohammed Ali in the first decade of the century tried to establish factories to supply the needs of his army. Spinning, weaving, the manufacture of glass, armaments and shipbuilding were introduced and directed by a system of monopoly, which were unable to survive after the downfall of his power. The industries sponsored by the Khedive Ismail about half a century later did not fare much better, though some traces of them survived. At the end of his time machinery and establishments were abandoned, but in some cases[7] — for instance, factories producing sugar — later developed further. The same was true of the industries which were built up during the First World War. The blocking of sea lanes created a favorable opportunity for production in the country, and factories for sugar, weaving, spinning, soap, furniture, iron, leather and perfume were established. With the return to normalcy and the opening up of transportation, most of the newly created industries could not stand the competition and had to be abandoned. Later, however,

7. Charles Roux and Henry Deherain, L'Historire de la Nation Egyptienne, VI. (Paris, 1936), p. 290.

the economic crisis, together with the nationalistic desire to make the country economically self-sufficient, served as incentives for reviving the old policies and making new efforts in the same direction.

The Bank Misr, founded with Egyptian capital in 1920, served, together with its affiliated societies, as the initiators of the movement for industrialization. Through this Bank the Government advanced considerable credit to manufacturers and artisans at interest rates from 4 to 6 per cent, rates which are regarded as very low in the Near East. The state also encouraged industrial development through an appropriate tariff policy. Since the early 1930's this policy has been extended.

Egypt, as a country of monoculture, is greatly dependent on the price of cotton in the world market. Prices were falling during the entire postwar period, reaching their lowest level at the beginning of the 1930's, when the world crisis developed. From over £E 7½ per cantar of cotton at the end of the war, the price fell to something over £E 4 in 1926–27 and decreased still further to £E 2 in 1931–32. Egyptian exports fell from £E 85.4 millions (in round figures) in 1920 to 31.9 in 1930 and 27.2 millions in 1931. The devaluation of her currency as a result of Great Britain's abandonment of the gold standard[8] made it even more difficult for Egypt to import needed manufactured goods. Hence, the Government adopted a strong protectionist policy, and encouraged production in local industries. Commercial treaties were denounced,[9] and high customs duties imposed on imports. The rate of customs duties, which amounted before 1930 to from 8 to 10 per cent of the value of the goods, rose to 23 per cent in 1932 and to an average of 40 per cent in 1939. Local industries were supported through cheap credits, preferences were given to local production in state purchases, and model workshops were founded to demonstrate methods. Since then various kinds of manufacturing have been developed in the country, notably spinning, haberdashery, silk and woolen goods, carpets, leather, furniture, soap, glassware, perfume, milk products and jams, and the exploitation

8. Egypt is part of the Sterling Bloc and her currency is linked with the British pound. The £E is worth approximately £ 1.06½ at par.

9. Before 1930 Egypt was bound by commercial treaties with many foreign states under the capitulatory regime to a limitation of the import tariffs to a general rate of 8 per cent *ad valorem* (with some exceptions) These treaties expired in 1930.

of mines and quarries was begun. Egypt has abundant raw materials for textiles (cotton and wool) and sugar cane. She has some manganese, the yearly output having reached 150,000 tons. She also has some oil (the 1939 output was 666,000 tons). The refineries belong to the Anglo-Egyptian Oilfield, Ltd. and the Government.

Egypt's industrialization, like Turkey's, has been promoted by the state. However, while in Turkey the large industrial enterprises are for the most part either state owned or belong to the state banks, industry in Egypt is mainly privately owned.

The most important branches of industry in Egypt are those which process agricultural products. There were at the end of the 1930's 145 establishments employed in cleaning cotton. Of these, nine belong to an Egyptian concern, and the rest to Britishers. Factories which extract oil from cotton seed are mainly financed by French capital. The production of cottonseed oil rose from 28,500 tons in 1933 to over 50,000 tons in 1937–38. French capital is also invested in sugar refineries.

The first establishments for sugar production were founded in the last quarter of the nineteenth century and further developed during and after the First World War. The volume, of sugar production reached 107,391 tons in 1930 and about 162,843 tons in 1939.

The production of textiles is a more recent development, initiated for the most part with Egyptian capital and actively supported by the Bank Misr. The first spinning enterprises of the Filatures Nationales d'Egypte were financed by foreign capital, but the Misr Spinning establishments founded at the end of the 1920's were financed through the Bank Misr and developed rapidly. They contain the largest industrial establishments in the Near East, have 60,000 spindles, and employ 20,000 workers. The production of cotton piece goods rose from a million square meters in 1930 to 159.5 millions in 1939.

The Bank Misr was also instrumental in establishing factories for the production of cotton yarn and silk. The cotton yarn production rose from 2,900 tons in 1930 to 25,800 tons in 1939. The production of leather and leather articles also developed in the interwar years, and with the help of government protection is supplying almost all the market in shoes. The production of cement

increased greatly, reaching 372,000 tons in 1939 as against 188,000 tons in 1930.

Besides the production of manufactured goods from local raw materials, an attempt was also made to manufacture goods using imported materials. To this group belongs the production of cigars, the tobacco being imported from Turkey and the Balkans. According to a statement of the Egyptian Minister of Commerce and Industry (in April, 1944), at the outbreak of the war Egyptian industries were able to meet the local demand to the extent of 100 per cent in sugar, alcohol, cigarettes, 99 per cent in lamp glass, 90 per cent in shoes, cement, soap, furniture, 80 per cent in matches and 40 per cent in cotton cloth. Some of these — sugar, alcohol, cigarettes, salt — were also being exported.

According to the census of 1937, industry in Egypt employs 489,027 workers, and together with building construction 609,735, which constitutes 8.2 per cent of the gainfully employed. The distribution of those employed in industry was as follows:

	%
Building	19.8
Clothing	17.7
Wood	12.6
Textiles	12.5
Metal	8.9
Food	7.8
Power and Electricity	6.6
Miscellaneous	14.1

As in the case of Turkey, the main branches of production (according to the number of persons engaged in them) were building construction, clothing and textiles. On the other hand, metal and food production employ relatively fewer workers in Egypt than in Turkey.

IRAQ

Iraq is less developed industrially than either Turkey or Egypt, although here, too, the Government has endeavored to encourage industrial production. The country lacks both manpower and sufficient capital for such development.

The steps taken by the Government to promote industrial production included the 1929 law to encourage industry (similar to the Turkish law of 1927), grants, reduction in taxes, custom duties,

credits through the newly founded Government Bank for Agriculture and Industry, and the introduction of uniforms for school children (1937) with the provision that the clothes were to be made from locally produced textiles.

Textiles were produced in Iraq before World War I in small artisan workshops. In the interwar years spinning establishments were founded which supplanted in part the local hand factories. The textiles produced in these factories supply the needs of the army, police forces, and school population, as do the tanneries, which were introduced in the last 15 or 20 years.

A postwar innovation is the tobacco factories, which have made fairly good progress, and the foundries, which were developed in connection with the erection of modern buildings. In 1938-39 there were several cigarette factories, with 780 workers,[1] and eight enterprises producing matches with 91 workers. Other enterprises worthy of mention are those producing wine, beer, etc., and flour mills. Soap, too, is produced in Iraq. Most of it comes from small handwork shops, but there are a few modern soap factories. Establishments for the production of cement blocks and tiles for bathrooms and roofs employ a larger number of workers.

MANDATED TERRITORIES

Syria-Lebanon and Palestine, as mandated territories, were, as we have seen, at a disadvantage, both because of the open-door clause which exposed them to foreign dumping and because of the reluctance of the mandatories to introduce innovations in the customs policy. The principle of tariff for revenue — the duties were levied *ad valorem* (12 per cent of the cost of the imported articles) — taken over from Ottoman times, was maintained by the mandatories, because customs constituted a significant part of the revenue.

To be sure, Palestine and Syria did not remain devoid of all protection against foreign imports. After 1927 the Government started to introduce in Palestine higher customs on some foreign goods which were also produced in the country. The Government also granted exemptions from duty on imported industrial machinery and raw materials. This protection came mostly too late,

1. The Iraq Directory, 1936, pp. 544–547 The oil fields, although bringing high royalties to the Government and employing large numbers of workers, are not included in this survey. The oil has been refined outside of the country.

however, or was insufficient. The institution of higher customs was always under pressure of the manufacturers or interested groups after the situation had become worse or the customs were not being handled properly. Both diminished the economic value of the new duties. For instance, at the end of 1936 specific duty rates not connected with the price on imports were introduced on bathtubs, but this was done only after the bathtub factory in Haifa had been compelled to close as a result of foreign competition.

Palestine had the lowest customs duties, not only in comparison with countries with high protective tariffs, but also in comparison with such a traditionally free trade country as Great Britain and the countries of the Near East. Syria fared somewhat better in this respect. She also benefited from the clause in the Palestine Mandate which provided that it may "conclude a special customs agreement with any state or territory which in 1914 was wholly included in Asiatic Turkey or Arabia" (Article 18). On this basis as early as 1921 a customs agreement was signed between Palestine and Syria along free trade lines. This agreement was renewed in later years with some small changes. Since Palestine had an expanding market on account of immigration and capital import, Syrian products found a good market there.

Syria and Lebanon are the home of old crafts. Before World War I these countries produced textiles, metal products, and jewelry. Silk, cotton and woolen cloth was woven and lace and carpets produced by artisans in their homes; only a limited amount of weaving was done in somewhat larger factories. The number of people employed in these enterprises at that time was estimated to constitute 10 to 15 per cent of the gainfully employed, and may have been even more.

During the war all production suffered a sharp decline. It is estimated that the silk spinning industry declined by 90 per cent.[2] There was also a decrease in the tanning industry. The downward movement of the traditional industries did not cease after the war, although for a few years an upward trend was to be noted.

In the postwar period the old industries had to cope with a number of factors. There was, as pointed out above, severe competition from other countries, while the Levant lost a great

2. A. Ruppin, Syrien als Wirtschaftsgebiet, pp. 131, 252; Abramowitz and Gelfat, Hameshek Haaravi (Tel Aviv, 1944), p. 158.

part of her markets through the destruction of the Ottoman Empire. Egypt, Anatolia (Turkey), and Iraq were protected from Syrian industries by tariffs. Only Palestine, because of commercial agreements, remained a market for Syrian products. At the same time, the purchasing power of the population in the Levant itself fell, both as a result of the declining tourist trade and of the low prices of agricultural products. The trend toward modernization, the changing mode of life, and the cheap imports of modern goods from other countries contributed toward the decline of the demand for the traditional consumer goods. The efforts of the Lebanese Government and the French High Commissioner to foster the revival of production of raw silk by arranging that the Levant should deliver it in considerable quantities to the silk textile industry of Lyon (France), had a temporary effect only, and after 1929 the downward trend began again. The old industries therefore were in an almost constant status of crisis and decline,[3] while some modern industries began to emerge. The latter were launched by emigrants returning to the country, who brought back their savings and experience, and received some support from the Government.

As early as 1924 industrial machinery was exempted from import duties, and two years later the tariff on certain imports was raised. In 1931 customs duties on machines and raw materials were abolished, while the duties on imported goods were increased. Since 1928 modern industrial activity has begun to manifest itself. The import of industrial machinery rose from 1,420 tons in 1925 to 3,159 in 1930, 3,830.7 in 1935 and 5,023.4 in 1937. With this machinery new silk and cotton spinning factories, cement, hosiery, knitting, cigarette factories, and flour mills were built. The economic crisis of the 1930's contributed to some extent to the development of new industries. Unemployment and falling prices of agricultural products created a large reservoir of very cheap labor, while the Government became interested in helping to create employment possibilities. Returning emigrants from the United States and French capital took advantage of these opportunities to found new enterprises.

According to a report from the year 1932, 148 modern factories

3. In Aleppo in 1913 there were 18,810 looms, at which 67,750 workers were employed; in 1926 there were only 5,000 looms, with 24,650 workers. S. B. Himadeh, ed. Economic Organization of Syria, p. 124.

already existed at that time, among them spinning and weaving mills, hosiery, match, cigarette and tobacco factories. In the next few years modern factories for weaving and for preserves (in Damascus) were also founded in Syria (industrial activity had at first been concentrated in Lebanon). A large weaving enterprise for cotton yarn was opened in Aleppo in 1936. The industrial development was somewhat slowed down in 1937–38 by the riots in Palestine, one of the main customers for Syrian goods.[4]

TABLE II

NUMBER EMPLOYED IN HANDICRAFTS, INDUSTRY, TRANSPORT AND BUILDING
IN SYRIA IN 1937

Industry	Number Employed
Motor and Lorry Drivers	12,000
Embroidery	3,400
Stockings, Socks and Knitting Factories	3,050
Mechanical Weaving	3,000
Cement and Concrete	2,120
Oil and Petrol	1,260
Harbor and Aerodrome Development	1,220
Boots and Shoes	1,100
Preserves and Sweetmeats	1,000
Various	5,000
Total	33,150

Source: Statistical Handbook of Middle Eastern Countries, Jerusalem, 1944, p. 104.

TABLE III

PRODUCTION OF CERTAIN COMMODITIES IN SYRIA AND LEBANON

	Syria		Lebanon	
	1935	1939	1935	1939
Macaroni	71 tons	1,095	750	1,375
Knit Goods	148,000 pieces ·	804,000	......	......
Socks and Stockings	2,424,000 pairs	2,988,000	250,000	230,000
Soap	1,610 tons	690	......	......
Shoes	1,200,000 pairs	750,000 (1938)	......	......
Chocolate	....	....	250 tons	235

Source: Statistical Handbook of Middle Eastern Countries, page 105.

According to an investigation made at the end of 1937, 33,150 individuals were employed in the new industries at that time.

4. Ministère des Affaires Étrangères, Rapport a la Société des Nations sur la Situation de la Syrie et du Liban, 1937, p. 15; 1938, p. 25.

These figures include 14,479 drivers and transport workers. If we deduct this group, there remain 18,670 industrial workers, which means a threefold growth since 1933, when the number of industrial workers amounted to only 5,800. The growth in output in the years 1932–38 was as follows: cement increased 5½ times; biscuits, 12 times; chocolate, 9½ times; cotton yarn, 4 times; stockings, 8 times.

While production in the new branches of industry was steadily increasing, the old branches were declining rapidly. In 1937 only 203,928 workers were employed in handicrafts, industry, and building in the main centers of the Levant, as compared with 309,526 in 1913. Of those employed in 1937, only 33,150 worked in the new industries (16 3 per cent), the others working either at home (40.1 per cent) or in small artisan shops (43.6 per cent). Over one-half of all the workers were employed in breeding silk-worms, silk and cotton weaving, and knitting; 5.4 per cent in shoemaking; 5 4 per cent in the production of food, 6 per cent as drivers, and smaller groups in clothing production and manufacture of cement, furniture, etc.

Palestine, too, experienced a decline in the interwar years in the traditional prewar handicraft production, but in this case the modernization of some of the old shops and the growth of new industries and factories have more than made up for the decline.

Before World War I Palestine had large wine cellars in Rishon leZion and Zichron Jacob, and some sizable soap factories in Nablus and Jaffa. The rest of the establishments — numbering, according to government figures, 1,235 — were mainly small homecrafts. Those included weaving of carpets, mats and cloth, potteries, flour mills, olive and sesame oil presses, production of glassware in Hebron and of religious souvenirs. In the postwar years some of these establishments in Palestine, like their counterparts in Syria, suffered as a result of changing market conditions, declining demand for relics, and the prohibitive policies of neighboring countries. (The soap industry, for example, suffered because of the tariff policy in Egypt.) At the same time other forces emerged which gave industry a new impetus.

The large-scale immigration of the postwar years has provided a market for industrial products and created a demand for consumer goods. Shelter, food, textiles, home furnishings, drugs, and many other commodities were needed for the new immigrants.

These products began to be produced in the country. Other factors favoring industrialization came with the immigrants: technical skill and capital.

Among the immigrants were technical experts and skilled workers. Capital was flowing into the country for entirely non-economic reasons. The situation in Germany and other Central European countries induced Jewish capitalists of those countries to seek a refuge for their funds in Palestine,· and their desire to invest this capital resulted in its use in financing young enterprises. Beginning with the immigration of the 1920's the country's economy, or at any rate that of the Jewish sector, began to develop in the direction of industrialization. At the time of the 1928 industrial census, taken by the Government in the first years of the new industrial trend, there were 3,505 establishments employing 17,995 persons, including owners. These establishments had a capital investment of £P 3,514,886 and an output valued at £P 3,886,149. They were mostly small artisan shops with an investment not exceeding £P 50 per unit. Almost one-third employed no hired workers; about 60 per cent employed from one to five workers each.[5] The total labor force amounted to 11,627 persons, over half of whom worked in establishments employing less than ten hired personnel. The Jewish share in these establishments is estimated as ranging from 35 to 50 per cent. This government survey was made when Jewish industrial activity was just beginning. The influx of immigrants, capital and skilled labor in the 1930's gave a greater impetus to young industry.

TABLE IV

ESTIMATED CAPITAL IMPORT INTO PALESTINE
(In £P 1,000,000)

1934	9.6
1935	12.0
1936	7.1
1937	5.1
1938	7.0
1939	7.0

A part of the new capital was used to import machines, which in turn served to build up industries, the value of machinery imports rising from £P 198,516 in 1925 to £P 448,707 in 1937. This

5. Said B. Himadeh, Economic Organization of Palestine, Beirut, 1938.

importation of industrial machinery was much heavier in Palestine
than in the neighboring countries. If the per capita import of
industrial machinery into Palestine in 1935 is taken as 100, the
corresponding figures will be: Egypt, 8.3, Syria, 7.9, and Iraq
(1934–35) 1.2.[6] Both the quantitative and the qualitative growth of
industry are revealed by the fact that the use of electric current for
industrial purposes rose from 2,190,464 kwh. in 1930 to 20,314,114
in 1937.

TABLE V

JEWISH IMMIGRATION AND THE GROWTH OF INDUSTRY

Year	Jewish Population	Industrial Establishments	Workers Employed in Them	Production	Capital Investments
					£P
1922	83,794	1,850	4,750	500,000	600,000
1937	395,845	5,606	30,040	9,060,000	12,700,000
INDEX NUMBERS					
1922	100	100	100	100	100
1937	472	303	632	1,812	2,117

Source: The Jewish Review, Vol. II, Nos. 2–3, July–October, 1944, p. 150

The connection between the increase of the Jewish population
and the growth of industry is shown by the data in Table V. The
Jewish population increased a little less than five times from 1922
to 1937, thereby creating a demand for commodities. The num-
ber of Jewish industrial establishments rose three times, and the
index numbers of their employees, the capital invested, and the
value of their production leaped from 100 to 632, 1,812, and
2,117, respectively.

Jewish industrialization served as a stimulus to non-Jewish
industrial activity, not as a competitor. Staffed with skilled and
semi-skilled Jewish workers, with a higher standard of living,
Jewish enterprise necessarily had higher production costs than
Arab enterprises employing cheap Arab labor. The higher-priced
Jewish articles could hardly compete with the cheaper Arab goods.
As a matter of fact, only ten per cent of the Jewish industrial
output is sold outside the Jewish market. At the same time, how-
ever, Jewish development served as an incentive for the Arabs.
Sometimes Jewish technicians were employed in Arab enterprises, or

6. League of Nations, Statistical Yearbook, 1935–36; 1936–37.

Arab workers got their first training in a Jewish firm. Although we have no full account of this development, the general trend may be gauged from the scanty material available.

According to the 1928 government census of industries, the number of Arab establishments was 1,023 before World War I, and those established in the postwar years up to 1928 numbered 1,373. The capital invested in these postwar enterprises was £P 613,000, as against 543,015 invested in 1914. The number of workers employed in these industries rose from 5,000 to 10,000. We have no exact information about the development since 1928. At the time of the 1931 census, 30,000 Arab wage earners were employed in industry and handicrafts. There are also some indications of the development during the 1930's. The so-called "traditional" Arab industries declined somewhat, while new ones grew up.

The production of soap, glass and mother-of-pearl goods declined or remained stationary. Arab soap production suffered the main decrease. The introduction of a high tariff on the import of soap to Egypt, the principal buyer of Palestinian soap, brought about a decline in this industry. The establishments in Jaffa, which produced the cheaper grades, were the hardest hit by this development; the higher priced grades manufactured at Nablus, not affected by competition, found a market. Similarly, Arab textile production — for the most part conducted on hand looms — suffered from Japanese competition in the 1930's. At the same time, certain other branches were developing. Of 522 oil presses in 1928, only 29 (six per cent) used mechanical power, while in 1941 as many as 144 establishments out of 605 (24 per cent) were mechanically driven.

Although no large modern industrial enterprises were established by Arabs in the interwar years, some fair sized factories were founded, which produced consumer goods. In some cases these factories were started by Arabs returning from America.

The production of tobacco and cigarettes developed in postwar years, as did factories for the production of cartons for the cigarettes. Factories to produce metals and matches were also founded by Arabs in this period. Most of these new establishments are technically modern and no longer of the artisan type. Arab small-scale industry and handicrafts expanded too, especially garages, mechanical repair shops, tailoring, cobbling and trades connected with building. As a result of the expanded building activity of

both the Jewish and Arab sectors, Arab industries developed in the interwar years which provided building materials (quarry, concrete, etc.). The old Arab stone buildings were replaced by modern concrete buildings. The Arab worker, previously unaware of concrete construction and building, learned all the occupations connected with this kind of building and found employment in Arab and Jewish building as well as in government large-scale construction schemes.

At the time of the 1931 census, industry and handicrafts employed 27,085 persons, or 12.6 per cent of the gainfully employed. On the basis of the growth of the Arab urban population in the 1930's (1931: 32.6 per cent; 1940: 35.0 per cent), it is estimated that the percentage of those employed in industry and handicrafts rose to about 14 per cent by 1940. The number of Arab workers in industry and handicrafts is estimated at 30,000. Among these groups of workers, however, are also included about 3,000 Arab workers who were employed in the so-called international concerns (refineries of the Iraq Petroleum Co., Shell Co., the British Spinney Co.) as well as in the match factory of Nur in Acco, and in the Dead Sea plants of Palestine Potash, Ltd.

DEVELOPMENTS DURING THE WAR

During the Second World War, industrialization in the countries of the Near East made great strides. Before the war these countries imported the bulk of their industrial goods, and when the sea lanes of the Mediterranean were blocked and shipping space became scarce, shortages of every kind of commodity resulted. The Allies were faced with the problem of supplying the local population as well as the needs of their armies. At the same time it became necessary to convert part of the local industries to the production of arms, munitions and other materials necessary to the winning of the war.

In 1940 Great Britain established the United Kingdom Commercial Corporation, Ltd., and in April, 1941, the Middle East Supply Center '(since 1942 an Anglo-American agency) was set up to conserve shipping space and to regulate and coördinate the flow to and production of necessities in this region. Both agencies endeavored to stimulate local production to meet the needs of both the civilian population and the armies. The Middle East Supply .

Center was instrumental in enlarging existing establishments or promoting new ones.[7]

Local government agencies as well as lend-lease machinery also did their part to stimulate the development of local production. What was to be produced and in what quantities was determined not by price but by consideration of strategic necessity. The rapid rise of prices created a favorable background for industrial (and agricultural) output.

This was also true in Turkey. Although the necessity for maintaining a large army and making heavy expenditures for defense diverted both labor and capital from industrial (and agricultural) production, the competitive buying of both belligerent groups helped to increase production. It is true that Germany could not fulfill the terms of the Turco-German commercial agreement of July, 1940, being unable to supply enough machinery and spare parts However, the United Kingdom Commercial Corporation and the British-American Coordinating Committee (set up in 1942) both bought Turkish products, and since the end of 1941 Turkey has also benefited from United States Lend-Lease. By coördinating the imports from the United Nations the aforementioned organizations helped Turkey to obtain some mining machinery, spare parts and industrial chemicals. In this way Turkish industrialization was able to proceed, at times slowly, after an initial drop in the years 1940–41.

The production of cement, for instance, dropped from 267,000 tons and 277,000 tons in 1940 and 1941, respectively, to 178,000 tons in 1942. Similarly, the output of glass and glassware dropped during the years 1942–43. The production of coal rose from 2.6 million tons in 1939 to 3 million in 1940 and 1941, dropped to 2.5 million in 1942 and rose again to 2.6 million tons in 1943 and 3.53 million tons in 1944. Sugar production dropped from 94,000 tons in 1939 to 57,000 tons in 1942, to rise again in 1943 to 96,420 tons. The production of chrome ore, as well as that of iron ore, declined steadily during the war years. On the other hand, the production of lignite (used as a substitute for coal), textiles (in spite of a slight set-back in 1942), woolen goods, and other commodities has steadily increased.

7 Francis Boardman, "Civilian Requirements from War to Peace: The Middle East Supply Center," The Department of State Bulletin, Vol. XIII, 339 (1945), pp. 994–99.

This increase in production was due to the reopening of smaller plants which had closed before the war, the completion of some plants which had been begun before the war, and the erection of some new ones. Among the new plants are a cotton mill at Malatya, equipped with 435 power looms and 26,400 spindles, the cement works at Sivas, a second paper mill and a cellulose factory at Izmir, a khaki dye plant, a sulphuric acid and superphosphate plant, a naphthalene factory at Karabuk, and a glue factory at Beykoz. All these factories are controlled by the Sumer Bank.[8]

The Ministry of Economy is engaged in working out a five-year plan for the promotion of Turkish industry, for which purpose 275 million Turkish pounds are to be allocated. The prime objects of the plan are to fill all existing gaps in Turkish factory activity and to erect woolen and spinning mills. Associated with it is a project for the establishment of a chemical and mechanical industry on a grand scale, foreshadowing an expenditure of 500 million pounds. It is also proposed that the factories and mines at present administered by the Sumer and Eti Banks should be transferred to seven special general Directorates.

The other countries of the Near East did not feel the strain of direct war preparations and mobilization to the same degree as Turkey, although Egypt was almost invaded by German forces, Syria was occupied by Free French forces, and British troops had to suppress the Rashid Ali revolt in Iraq in 1941, while Palestine supplied a considerable number of volunteers to the British forces. But the cutting off of the sea lines after Italy's entry into the war brought about two contradictory trends: one which acted as a detriment to the expansion of the economy; the other, and more important, which stimulated production. The Allies encouraged local agricultural and industrial production and endeavored to aid their development. At the same time, the needs of the Allied armies and the local population, which could no longer be supplied by imports, served as an incentive to increase production. There was the problem of fuel for industrial uses: in Egypt the lack of imported coal was solved by using oil; in Palestine electricity was used instead.

Experiments were tried with new crops which could be used for industrial reprocessing (flax in Egypt), which however were not

8. The Malatya cotton mills are controlled by the Sumer Bank in conjunction with the Ish Bank and the Agricultural Bank.

always successful. Despite the efforts of the United Kingdom Flax Controller in Egypt, for instance, and the initial enlargement of the area under flax cultivation (10,000 acres in 1939 to 51,500 acres in 1941), the quality of the fibre produced proved disappointing, and the cultivation has fallen back (5,000 acres in 1944). The Middle East Supply Center was instrumental in Egypt in the development of production of tungsten, tin, and the mining of chrome ore (to make bichromates and chromic acid for use in dyeing textiles and tanning leather).

Machinery which could no longer be used in the production of silk, because of the shortage of raw materials, was converted to the production of cotton goods. The cotton piece goods production rose from 77 million square yards in 1937 to 370 in 1942. The production of sugar rose from 160,000 tons in 1937 to 190,000 tons in 1943, and that of beer increased from 1½ to 9 million gallons. The production of paper, glassware, soap, alcohol, glycerine, paints, cement, jams and some chemicals increased. Among the new industries were dehydrated vegetables, glucose, lead refining, cast iron pipes, steel castings, lead tubes, copper sulphate and cigarette lighters. The production of spare machine parts was also started, as well as of simple electrical appliances such as batteries.[9]

In Palestine the growth of industry came mainly after 1941. In the earlier years of the war, most of the requirements of the British Army were ordered outside of Palestine; military contracts placed with local manufacturers did not exceed £P one million (about four million dollars) in 1940. In 1941 the figure was £P 4,000,000 and in 1942 about £P 8,000,000, a figure almost equal to the total prewar industrial output of Palestine. The demands of the military plus the necessity of supplying the civil population with consumer goods served as incentives to the expansion of industry. The number of Jewish industrial establishments rose from 5,602 in 1937 to 6,116 in 1943, the number of gainfully employed in industry from 28,616 to 52,000 and the value of the output from £P 8,526,000 ($34,104,000) to £P 38,000,000 ($152,000,000).

This expansion was effected for the most part by the utilization of existing machinery, plants and technical skill, which had not been utilized to the highest capacity before the war. In the years

9. Cf. Table IV (Wartime Development of the Capacity of Egyptian Industry) printed as Appendix IV in Egypt, Review of Conditions, May 1945, London, Department of Overseas Trade, pp. 37–40.

1939–42 about 500 new Jewish establishments were organized (a growth of some 10 per cent), while there was a 70 per cent growth in the number of gainfully employed in industry, and the value of the output increased by 300 per cent. In 1937 there were 16 Jewish establishments which employed more than 100 persons; in 1942, 50 such enterprises were in operation. The main expansion took place in the production of textiles, metal and machinery, chemicals and foodstuffs, the largest gains being in food processing and in chemicals. On the other hand, there has been a standstill, or a contraction, in the production of wood, printing, paper, stone and electrical equipment. This is attributable to the lack of raw materials (paper, printing) and the stoppage of building.

Many of the new undertakings, as well as some of the old ones which have been enlarged, have added a host of new articles which had previously not been manufactured in Palestine; for example, industrial machinery and instruments, motor-car parts, textiles, domestic and kitchen utensils, agricultural machinery and tools, pharmaceuticals, window glass, safety razors and blades, carbonic acid, oxygen, sulphuric acid, brass and steel castings, precision instruments, electric batteries and plastics. A new industry is the diamond cutting and polishing industry, which grew up during the war and now employs over 4,000 workers. Production grew from 50,000 karats in 1943 to 130,000 in 1945. Most of these products have begun to be exported.

Petroleum production and that of chemicals from the Dead Sea were greatly expanded. Before the war the oil coming from Iraq by pipeline was sent to other countries in a crude state. In 1939 the Iraq Petroleum Co. began erection of the refinery in Haifa, which was finished during the war. In 1939, 92,000 tons of crude oil were refined, and this increased to 1,292,000 tons in 1941 and 3,363,000 in 1944. Today the refineries not only supply the local market with petroleum, benzine, diesel motor oil, etc., but produce a surplus for export, which in the last years amounted to some £P 5,000,000 ($20,000,000) annually. The production of potash and bromine by Palestine Potash, Ltd. is said to have more than doubled.

The Arab industrial sector in Palestine also made some gains during the war. In the years immediately preceding the war, Arab economy suffered a decline, the years of unrest and terror having undermined its strength. There was a decline in the textile

industry, production of soap, arts ouvenirs, etc., the great decline being in the Arab cigarette factories, where production fell from 540 tons in 1935 to 280 tons in 1939. The war brought about the end of the riots and the emergence of Allied armies as customers for goods and services, creating a favorable background for development. Since the Arab industry consisted mostly of smaller shops, which were unable to undertake larger orders, the Government had to supply them with raw materials and loans. Such assistance was extended to soap and textile factories. According to the government censuses of 1940 and 1943, the number of Arab establishments grew from 350 in 1939 to 1,558 in 1942 and the personnel employed from about 4,000 to 8,804.[1] In the textile industry, mainly in Gaza and Migdal, the number of looms rose from 400 (1939) to 1,500. During the war years several modern textile factories, employing 50–60 workers each, were founded and mechanical looms were introduced. In 1942 a weaving plant was founded in Nablus which employs 100 workers; in 1944 a glass factory was opened in Jaffa. Cigarette production grew from 450 tons in 1940 to 700 in 1942, with further increases in the following years. New Arab plants for tanning, the production of jam, soap, etc. were also established.

The success of wartime industrial development depended on the availability of machinery, or the possibility of importing it, and on the existence of technically trained personnel. Both were more available in Egypt and Palestine than Iraq or Syria. In Iraq the only new industries which sprang up during the war were the manufacture of hand tools from scrap metal and a mill for the crushing of oil seeds. Of the prewar industries, tanning and glass production made some gains, while the production of textiles, boots and shoes remained the same. The production of soap even declined somewhat, because of a lack of the necessary oils.

The Levant made more progress. Syria and Lebanon were aided by the Emergency Economic Office established in 1942 by General Cartaux. Another impetus to local manufacture came from the changes in trade caused by the war. Before the war, for instance, Syrian wool was exported to America and Europe, and cheap woolen goods were imported.[2] During the war, because of

---

1. Government of Palestine, Department of Statistics, General Monthly Bulletin of Current Statistics, April, 1945, pp. 178 ff.

2. On the eve of the war, wool exports amounted to 1,800 tons annually and imports of woolen goods to 2,100 tons.

restricted imports, local wool was used for textile manufacturing. The same was true of soap. Before the war the production of soap was limited to local requirements, the bulk of the crude olive oils being sent to Italy. Recently two large soap factories were established. The manufacture of cotton goods, silk and silk goods was stepped up, the tanning and leather industry made some gains, getting their raw materials from local sheep and cattle breeders. In Syria the textile and glass industries expanded. In Lebanon 140 new establishments, mostly small plants, were founded, as well as a sugar refinery and plants for the production of oil, soaps, alcohol and dyes. Factories for tinned fruit and vegetables — two in Syria and two in Lebanon — which sprang up to meet military requirements, are a new development. In 1942, Syrian factories produced 6,000 tons. The production of cement and clay also increased rapidly during the war years. This growth in industrial output came about in the later years of the war, almost every branch of production having declined in the years 1939-41.

POSTWAR PROBLEMS

When the war ended, the countries of the Near East were faced with the difficult task of converting industries from a war footing to civilian peacetime production. At first, it is true, the falling off of Army orders had only a slight influence on the industries of these countries, since the actual production of arms played only a minor rôle in these countries. Most of their enterprises were either producing spare parts or consumer goods, which were in part bought up by the Army. The technical task of reconversion was therefore an easy one, and in the face of the immense lack of consumer goods, the products readily found a market. Some of the new industries, however, will be unable to survive: those which could exist only as long as no other source of supply was available. On the other hand, some of them are basically sound, despite the fact that the new establishments here have required higher investments than similar establishments in other countries. In Egypt there are factories which are satisfactorily equipped technically. According to British opinion on Palestine "the clothing and fashion trades are typically Jewish industries in many countries. Palestine will in all probability be in a position to continue this industry ... the same applies to many branches of the fancy leather industry ... the future of the diamond-cutting industry, too,

seems assured. There is a good potential future for the plastic industry ..." while the "pharmaceutical industry is partly sound. ..."[3] A number of new industries in other Near Eastern countries have equally favorable prospects.

Nevertheless, all these countries are faced with the problem of maintaining their present scale of industrial activity. Germany, Italy and Japan, to be sure, are no longer competitors, and cannot dump their products on Near Eastern markets; on the other hand, the war has brought about many changes: first of all, the question of prices. In the whole of the Near East there is a clash between the prevailing price levels and those in the United States and Great Britain, and there is the problem of adaptation to the world market and normalization of prices.

The increased spending of the Armies and the closing of the sea lanes during the war resulted in a sharp rise in prices. The wholesale price index for Egypt (January–June, 1939 = 100) moved to 213 in 1942, 311 in 1944, and 325 in May, 1945. The wholesale price index for Palestine rose to 260 in 1942; 325 in 1944, and 333 in May, 1945. The rise has been still steeper in Iraq, where the general wholesale price index rose to 534 in 1943. In Syria, at that time, it was 953; in Turkey, 446. Actual prices at a given time may be even higher than the price indices indicate. Some commodities figuring in indices may either be unavailable or to be found only in the black market. A ten-cent bar of soap, for instance, cost a dollar in Lebanon in January, 1946. Shoes were from $45 to $74 a pair, and $35 was paid for an aluminum pan.[4] These rising prices have shifted the countries of the Near East from among those with the lowest cost of living and of production and placed them among those with the highest in the world.

The productivity of labor in the Near East reaches only one-third to one-half of American or British standards. Rationalization and increased productivity of the worker, about which there is much discussion in Palestine and Egypt, is a slow process. In the Near East increased local production and, to a certain extent, even the maintenance of existing production depend mainly on imports of machinery and capital goods, which are curtailed by the shortage throughout the world and still more by the general financial

3. Palestine. Review of Commercial Conditions, London, 1945.
4. Christian Science Monitor, January 21, 1946.

situation.[5] During the war years a great deal of the existing machinery and tools have worn out, and such equipment is not easy to get anywhere nowadays. Moreover, the Near Eastern countries, either belonging to the Sterling Bloc or having a monetary agreement with Great Britain, are unable to trade with countries like the United States or Switzerland for lack of "hard currency," and the allotments of the pool in London are very small. Only Iraq's allotment, according to the Anglo-Iraq Scarce Currencies Agreement for 1945, was higher than her earnings in dollars.

Turkey does not belong to the Sterling Bloc and has a financial agreement with Great Britain which provides for a kind of clearing. According to this Trade and Payment Agreement, the Central Bank of the Turkish Republic is to open a "Turkish account" at the Bank of England into which all payments due to Turkey are to be paid. The amount which accumulates in the "Turkish accounts" can be used for payments inside the Sterling area and 20 per cent also for payments outside the Sterling area.

In the second place, a great part of the resources of this region are frozen in London as so-called "war balances."[6] Great Britain is neither able to pay her debts in cash now nor to deliver machinery in payment. Neither is she interested in doing this, since she must export 150 per cent as much as she did in prewar years if she desires to restore her prewar standard of living. Since VE and VJ Days she has succeeded in greatly increasing the exports of hardware, tools, electrical equipment, glassware, pottery, and artificial silk. She is lagging behind with textiles and other goods, partly because during the war local production has sprung up in some countries which were formerly customers for British textiles. British circles are therefore very reluctant to deliver machinery to the newly-created industries of the Near East in order to avoid strengthening competitors.

The United States loan to Great Britain and the consequent easing of currency restrictions will, of course, improve the situation for the Near Eastern countries, but the repayment of the so-

5 Cf also, B D. Weinryb, "Britain's War Debts in the Middle East," *Asia and the Americas,* February, 1945.

6 The importance of these war balances could be gauged from the fact that 120 million pounds which Great Britain owes to Palestine equals almost the whole import of capital during the interwar years, and the 350 million pounds which Great Britain owes to Egypt equals about ten years of government revenue according to prewar situations.

called "war balances" remains unsettled. However, the freeing of future dollar earnings of these countries, which Great Britain has promised, will bring partial relief. Egypt is likely to continue to experience a deficit in dollars. The principal export commodity of Egypt is long-staple cotton, and its import into the United States encounters the opposition of the protective tariff and the competition of rayon and nylon. Egypt's dollar earnings are estimated at 13–14 million, while her import needs from America amount to about 50–60 million. Iraq will probably also not have an ample dollar income, although the improvements which were made last year in the date-packing plants at Basra meet the sanitary regulations of the United States and the Iraqi dates will now find a market here. Raw wool, hides, and skins from Iraq will also find a good market in the United States.

These dollar earnings, however, do not seem to suffice for the payment of necessary imports. The decrease of American exports via Iraq and Iran to Russia has brought about a decline in Iraq's dollar earnings, which are put at four million dollars, yet she needs 14–15 million dollars to pay for her imports. Turkey finds herself in a peculiar position. Great Britain is reluctant to buy in Turkey at the current high prices and thus enable Turkey to accumulate balances in the "Turkish accounts." Great Britain would have to balance her purchases in Turkey with the delivery of other commodities plus a certain percentage in hard currencies.[7]

Palestine alone has, and is likely to continue having, a surplus of dollars originating from exports and from invisible items. The exports from Palestine to the United States grew during the war years, mainly because of the development of the diamond industry, from £P 97,000 ($48,500) yearly average in 1937–39 to £P 2,339,000 ($9,358,000) in 1943. Since 1943, the diamond exports from Palestine to the United States have almost doubled. A considerable flow of dollars to Palestine also results from institutional contributions and private investments. In the years 1939–43 institutional remittances from the United States averaged £P 1,853,000 ($7,412,000). In 1944 and 1945 they were considerably higher, and those for 1946 and 1947 will undoubtedly be still higher, as a result of the U.J.A. drives.

7. In the last months of 1946 a solution was sought in the devaluation of the Turkish currency. It remains to be seen, however, to what extent the devaluation of the currency will influence Turkish exports in the long run.

Near Eastern industries are also faced with other problems. Some of the new factories used raw materials which had to be procured from distant places, and they received the assistance of the Middle East Supply Center in this connection. Now, however, the situation is different. The textile industry in Palestine, for instance, depends upon the import of yarn which comes mostly from India, and upon local production. There are in Palestine two factories for production of yarn with about 15,000 spindles. In order to meet the textile industry's requirements for yarn, the number of spindles would have to be increased fourfold. However, influential circles in England are opposed to exporting textile machinery, since they want to retain foreign markets for their cloth. As a result textile production had to be curtailed, although lately yarn has been imported from Brazil, and is just beginning to come in from Italy.

For Palestine another problem exists, that of the Arab boycott. During the war Palestine's industry started to produce not only for the local market but for export to neighboring countries as well; pharmaceuticals, clothing, shoes, dyes, glassware found a market in the Near East. In the last year, the agitation in some of the Arab countries, mainly in Egypt, against the products of Jewish industry in Palestine has become more pronounced. In conjunction with its political agitation, the Arab League launched a boycott of "Zionist" industry at the beginning of 1946. Palestine, it is true, imports much more from neighboring Arab countries than she exports to them (£P 8,000,000 against £P 3,250,000 in 1944) and is therefore in a position to retaliate against the boycott; but in the long run the economic policy of the Arab countries may nevertheless pose a serious problem for Palestine.

All these problems are of grave concern to the countries in question. In almost all of the Near Eastern countries there are discussions about maintaining industries, reducing prices, increasing productivity, and re-equipping factories with the most up-to-date tools. They regard it as necessary to devote more attention to technical education, and a demand is being voiced for the establishment of state industries or of state-supported industries. In the face of the necessity for enlarging her armies and increasing her defense forces in view of the British withdrawal, there is also in Egypt a demand for the creation of new factories for arms production. And the introduction of high customs tariffs, to

protect local industries from foreign competition, is advocated.

It is difficult at this point to foresee the course of events in the Near East and to predict which war industries will survive in the postwar period. This will depend not only on economic factors, but on political factors as well. One thing seems to be clear, however: the war gave industrialization a great impetus. Whatever the lot of the individual factories, industry is sure to occupy a prominent place in the economies of these countries, or at least in Egypt, Palestine and Turkey. Moreover, the Near East is becoming a center of oil production and refining, and the availability of fuel will undoubtedly influence industrialization. These facts are of importance, both for the future situation of the Near East and the trade relations with that region.

<div align="right">

BERNARD D. WEINRYB.

</div>

NEW YORK CITY

RECENT PUBLICATIONS

ABELARDE, PEDRO E. American Tariff Policy Towards the Phillipines, 1898–1946. New York: King's Crown Press. 1947. pp. 233. $2.75.

ABRAMS, CHARLES. A Housing Program for America. New York: League for Industrial Democracy. 1947. pp. 32. 25 cents.

AMONN, A., ET AL. Konkurrenz und Planwirtschaft. Bern, Switzerland: Verlag A. Francke. 1946. pp. 240. s.fr. 19.

ARTZ, F. B. L'Enseignement Technique en France pendant l'Epoque Revolutionnaire, 1789–1815. Paris: 1946. pp. 53. [Reprinted from the "Revue Historique."]

AUDIT UNION OF THE WORKERS' AGRICULTURAL COOPERATIVE SOCIETIES, LTD. The Palestine Agricultural Economy under War Conditions. Tel Aviv, Palestine: The Union. 1946. pp. 51.

BACKMAN, JULES AND GAINSBRUGH, MARTIN R. Profits in the National Economy. New York: National Industrial Conference Board, Inc. 1947. pp. 32.

BELGIUM, MINISTÈRE DES AFFAIRES ÉCONOMIQUES. Démographie De La Belgique de 1921 a 1939. Brussels: Office Central de Statistique. 1943. pp. 460.

——. Bulletin de Statistique. Brussels: Office Central de Statistique. January 41 issues and later.

——. Bulletin de Statistique. Brussels: Office Central de Statistique. [January 41 issues and later.]

——. Annuaire Statistique de La Belgique et du Congo Belge. Brussels: Office Central de Statistique. Vol. LXII, 20 francs, 1940; Vol. LXIII, 20 francs, 1941; Vol. LXIV, 40 francs, 1942; Vol. LXV, 20 francs, 1943; Vol. LXVI, 20 francs, 1944.

——. Recensement Economique et Social, Les Etablissements Commerciaux. Brussels: Imprimerie fr. Van Muysewinkel. February, 1937. 5 volumes.

BELGIUM, MINISTÈRE DE L'INTERIEUR. Economische en Sociale Telling. Brussels: Office Central de Statistique. February, 1937. 2 volumes.

BUREAU OF THE BUDGET. The United States at War. Washington, D. C.: War Records Section of the Bureau. 1947. pp. 555. $1.00. [Prepared under the auspices of the Committee of Records of War Administration by the War Records Section, Bureau of the Budget.]

CALIFORNIA, STATE RECONSTRUCTION AND REEMPLOYMENT COMMISSION. The Steel and Steel-Using Industries of California. Sacramento, California: The Commission. 1946. pp. 408. $1.50 plus tax.

CHANDLER, LESTER V. A Preface to Economics. New York: Harper and Brothers. 1947. pp. 289.

CIVILIAN PRODUCTION ADMINISTRATION. Industry and Labor Advisory Committees in the National Defense Advisory Commission and the Office of Production Management, May 1940 to January 1942. Washington, D. C.: Bureau of Demobilization. 1947. pp. 233.

CLOUGH, SHEPARD B., AND COLE, CHARLES W. Economic History of Europe. Boston, Massachusetts: D. C. Heath and Company. 1946. pp. 904. $5.00.

COMMITTEE ON POSTWAR TAX POLICY. A Tax Program for A Solvent America. New York: The Committee. 1947. pp. 170.

DAGGETT, STUART AND CARTER, JOHN P. The Structure of Transcontinental Railroad Rates. Los Angeles, California: University of California Press. 1947. pp. 161. $4.00.

DAVIS, JOSEPH S. International Commodity Agreements: Hope, Illusion, or Menace? New York: Committee on International Economic Policy. 1947. pp. 81.

DUBLIN, LOUIS I., AND LOTKA, ALFRED J. The Money Value of a Man. New York: The Ronald Press Company. 1946. pp. 214. $6.00.

EDITORS OF LOOK. Oil for Victory. New York: Whittlesey House (McGraw Hill). 1946. pp. 287. $3.50.

FESLER, JAMES W. Pulp and Paper Policies of the War Production Board and Predecessor Agencies, May 1940 to January, 1944. Washington, D. C.: Civilian Production Administration, Bureau of Demobilization. 1946. pp. 161. [Special Study No. 7, Historical Reports on War Administration: War Production Board.]

FREHNER, EDWIN. Die Reserven in der Unternehmung. Zurich, Switzerland: Schulthess and Company. 1946. pp. 151.

GEIJER, ERLAND. How Swedish Tax Laws Affect Persons Removing into or from Sweden or Residing Abroad. Sweden: Svenska Handelsbanken. 1946. pp. 55.

GEMMILL, PAUL F., AND BLODGETT, RALPH H. Current Economic Problems. New York: Harper and Brothers. 1947. pp. 760. $4.00. [Third Edition.]

GOPAL, M. H. Basis of Public Salaries. Mysore, India: The University of Mysore. 1946. pp. 140.

——. The Structure of a Permanent Excess Profits Tax. Mysore, India: The University of Mysore. 1946. pp. 51.

HANSEN, ALVIN H. Economic Policy and Full Employment. New York: McGraw-Hill Book Company. 1947. pp. 340.

INTERNATIONAL LABOUR OFFICE. Public Investment and Full Employment. Montreal, Canada: The Office. 1946. pp. 348. $2.25.

ISERMAN, THEODORE R. Industrial Peace and the Wagner Act. New York: McGraw-Hill Book Company. 1947. pp. 91. $1.50.

KENT, RAYMOND P. Money and Banking. New York: Rinehart and Company, 1947. pp. 702.

KENTUCKY, DEPARTMENT OF REVENUE. Income Tax Law and Regulations. Frankfort, Kentucky: The Department. 1946. pp. 186.

KRISTENSSON, FOLKE. Studier I, Svenska Textila Industriers Struktur. Stockholm, Sweden: Industriens Utredningsinstitut. 1946. pp. 503. [Including an English Summary.]

LEAGUE OF NATIONS. Course and Control of Inflation. New York: Columbia University Press. 1946. pp. 136. $1.25.

LUNDBERG, GEORGE A. Can Science Save Us? New York: Longmans, Green and Company. 1947. pp. 122. $1.00.

MACFARLAND, GEORGE A., AND AYARS, ROBERT D. Accounting Fundamentals. New York: McGraw-Hill Book Company. 1947. pp. 759. $4.50. [Second Edition.]

MACIVER, ROBERT M. The Web of Government. New York: The MacMillan Company. 1946. pp. 498. $4.50.

METZ, HAROLD W., AND JACOBSTEIN, MEYER. A National Labor Policy. Washington, D. C.: The Brookings Institution. 1947. pp. 164.

MILLER, RAYMOND W. The Corporation — A Brotherhood of Service. Ottawa, Canada: World Trade Relations, Ltd. 1947. pp. 19. $1.00.

MINISTERO PER LA COSTITUENTE. Rapporto della Commissione Economica Presentato all'Assemblea Costituente. Roma, Italy: Istituto Poligrafico dello Stato. 1946. [II Industria, III Problemi Monetari e Commercio Estero, IV Credito e Assicurazione, V Finanza, VI Appendix.]

MOORE, GEORGE ALBERT. Jean Bodin's Response (1568–78) to Malestroit's Paradoxes (1566). Chevy Chase, Maryland: Country Dollar Press. 1946. pp. 90. [An English Translation.]

PATTERSON, ERNEST MINOR. An Introduction to World Economics. New York: The MacMillan Company. 1947. pp. 704. $5.00.

PEARN, B. R. The Mixed Races of Burma. Malvern, England: Le Play House Press. 1946. pp. 8. [Racial Relations Studies in Conflict and Coöperation No. 3.]

——. The Indian in Burma. Malvern, England: Le Play House Press. 1946. pp. 36. [Racial Relations Studies in Conflict and Coöperation No. 4.]

PELZEL, SOPHIA ROGOSKI. American Intervention in Siberia, 1918–1920. Philadelphia, Pennsylvania: University of Pennsylvania. 1946. pp. 98.

PERROUX, FRANCOIS, ED. Les Caractères Contemporains du Salaire. Paris: Presses Universitaires de France. 1946. pp. 160. [Publication of the Institut de Science Économique Appliquée.]

——. Salaire et Rendement. Paris: Presses Universitaires de France. 1946. pp. 232. Publication of the Institut de Science Économique Appliquée.

PETRUZZELLI, NICHOLAS MICHAEL. Some Technical Aspects of Foreign Trade Statistics with Special Reference to Valuation. Washington, D. C.: The Catholic University of America Press. 1946. pp. 252.

POSTHUMUS, N. W. Inquiry into the History of Prices in Holland. Leiden, The Netherlands: E. J. Brill, Ltd. 1946. pp. 660. 50 guilders.

PRATHER, CHARLES L. Money and Banking. Chicago, Illinois: Richard D. Irwin. 1946. pp. 803. $5.00.

PRIME, JOHN H. Investment Analysis. New York: Prentice-Hall, Inc. 1946. pp. 442. $4.65.

PRITCHARD, L. J. Kansas Banking During the War Economy Period 1939–1945. Lawrence, Kansas: University of Kansas Publications. 1946. pp. 86 [Industrial Research Series No. 7.]

SHAPIRO, ELI. Credit Union Development in Wisconsin. New York: Columbia University Press. 1947. pp. 174. $2.50.

TAX INSTITUTE. Capital Gains Taxation. New York: The Institute. 1946. pp. 106. [A Panel Discussion by: Thomas N. Tarleau, Chairman; Franklin Cole; Harold Groves; George O. May; Harry Rudick; Beardsley Ruml; Eustace Seligman; Lawrence Seltzer; Carl Shoup.]

TWITCHELL, K. S. Saudi Arabia. Princeton, New Jersey: Princeton University Press. 1947. pp. 192. $2.50.

VINCE, ANDRÉ. La Doctrine Americaine de la Porte Ouverte en Chine et son Application dans la Politique Internationale. Geneva, Switzerland: L'Institut Universitaire de Hautes Études Internationales. 1946. pp. 148.

WEDGEWOOD, SIR RALPH L. International Rail Transport. New York: Oxford University Press. 1947. pp. 162. $3.00.

WRIGHT, DAVID McCORD. Economics of Disturbance. New York: The MacMillan Company. 1947. pp. 115. $2.50.

PRICES AND WAGES UNDER BILATERAL MONOPOLY[1]

SUMMARY

Present state of the theory, 503. — Three cases, 505. — Diagrammatic presentation, 507. — Extension to labor-management relations: assumption I, 509; assumption II, 510; assumption III, 511. — Some complicating factors: supply of labor, 516; efficiency of labor, 518; discontinuities in the value product functions, 519. — Wage theory on the general equilibrium level, 520. — Reconsideration of the theory of bilateral monopoly, 523. — Institutional obstacles to movement to the contract curve, 524. — The contract curve and the concept of the "solution," 528. — The range of indeterminateness, 530.

The theory of bilateral monopoly has a direct bearing, not merely on price and output in certain product markets, but also on the more generally significant problem of wage rates and employment under collective bargaining. In the contemporary institutional setting, it is impossible to develop a reasonably realistic value theory without relying heavily on the theory of bilateral monopoly. It therefore seems worth while to examine the present state of bilateral monopoly theory and the possibilities of adapting it to the conditions prevailing in the labor market. What are the "accepted views" concerning the determination of price and output in bilateral monopoly? Should these be regarded as valid or should they be modified and supplemented in certain respects? How could they be adapted to fit the conditions in which wage rates and employment are determined by collective bargaining? The following discussion centers on these questions.

THE PRESENT STATE OF THE THEORY ·

The present stage of development of the theory of bilateral monopoly is well described in a note by Gerhard Tintner. Tintner's

1. This article developed from detailed discussions with Professor Howard S. Ellis. Much of what is contained in the article is his contribution.

note is based mainly on Professor Bowley's position, as expressed in 1928 under the influence of Wicksell's and Schumpeter's criticism of his earlier opinions. Professor Hicks has further elaborated on Bowley's revised views.[2] Essential features of this position are also accepted in Heinrich von Stackelberg's presentation of the problem.[3] No attempt will be made to trace the line of descent of this position beyond what was just said.

A brief summary of these "accepted views" will now be presented. The simplifying assumptions used in the following account are not precisely identical with those adopted by previous writers, but in no essential respect will the account given here deviate from previous statements of the theory.

Let us assume that one of two firms produces a raw material which is used by the other firm to manufacture a finished commodity. The producer of the finished commodity is a monopsonistic buyer of the raw material. He may be a monopolistic or a competitive seller of his own product. At any rate, his customers are competitive buyers, and the industries from which the producer of the raw material buys *his* materials are competitive industries. Consequently, the producer of the finished commodity is faced with the usual kind of demand function for his product, and the producer of the raw material is faced with the usual kind of supply function for the materials *he* uses. The "unusual" problems — those typical of bilateral monopoly relations — arise in connection with the passing of the raw material in question from its producer to the producer of the finished good. The questions to be answered are these: what quantity of the raw material will be produced and sold to the manufacturer of the finished commodity; and what price will be paid for the raw material? All other price and quantity questions arising in the model are answered by the traditional theories of monopoly and competition.

2. Cf. Joseph Schumpeter, "Zur folgenden Arbeit Knut Wicksell's," Archiv fuer Sozialwissenschaft und Sozialpolitik, Vol. LVIII, 1927, pp. 250ff.; Knut Wicksell, "Mathematische Nationaloekonomie," ibid. pp. 275ff.; A. L. Bowley, "On Bilateral Monopoly," Economic Journal, December, 1928, pp. 651ff.; J. R. Hicks, "Annual Survey of Economic Theory: The Theory of Monopoly," Econometrica, January, 1935, pp. 16ff.; Gerhard Tintner, "Note on Bilateral Monopoly," Journal of Political Economy, April, 1939, pp. 263ff. Bowley's earlier position was presented in Mathematical Groundwork of Economics, Oxford, 1924, p. 62.

3. Heinrich von Stackelberg, Marktform und Gleichgewicht, Wien und Berlin, 1934, pp. 24ff.

According to what we have called the "accepted views," the answers to these questions are as follows. Three cases should be distinguished, (*a*) the dominance of the buyer of the raw material (that is, of the producer of the finished commodity), (*b*) the dominance of the seller of the raw material (that is, of its producer), and (*c*) maximization of the joint profit of these two producers.

Case (*a*) is defined by the characteristic that the seller of the raw material is sufficiently weak in relation to the buyer to set up a supply schedule, which coincides precisely with his *marginal cost function* (*MC*). The buyer, who in this case "dominates," chooses the point along this function which is most favorable to him. This he will do by equating the marginal function corresponding to the supply curve, that is, the function which is *marginal to the marginal cost function of the seller* (*MMC*),[4] with the net marginal value product function (MVP) of the raw material in his own enterprise. The marginal function to the marginal cost function of the seller (MMC) is the marginal cost of the raw material to the buyer.[5] He maximizes his profit by equating this with the marginal value product (MVP) of the raw material to him. The price paid for the raw material will be that indicated by the supply schedule of the seller, that is, by his marginal cost function (MC) for the quantity equating MMC with MVP. These conditions, as well as those characteristic of cases (*b*) and (*c*) may be read from Figure I.

Case (*b*) is defined by the property that the buyer of the raw material — that is, the producer of the finished commodity — is sufficiently weak in relation to the seller to set up a demand function for the raw material. This demand function is the (net) *marginal value product function* (MVP) of the raw material to him. The seller of the raw material, who in this case "dominates," chooses the point along this function which is optimal to him. This he finds by equating the marginal function corresponding to the demand curve of the buyer, that is, the function which is *marginal to the marginal value product function* of the raw material to the

4. In other words, the function bears the same relation to the marginal cost function as the latter to the average cost function. If the aggregate cost for any quantity x is $\phi(x)$, then the average cost is $\dfrac{\phi(x)}{x}$, and $MC = \phi'(x)$, $MMC = \phi'(x) + x\phi''(x)$.

5. For the marginal cost function itself is the supply function of the seller and, thereby, the average cost function to the buyer.

buyer (MMVP),[6] with his own marginal cost function (MC). The reasons are analogous to those discussed in connection with Case (*a*). The price charged will be that indicated by the demand function of the buyer (MVP) for the quantity equating MMVP with MC.

Cases (*a*) and (*b*) are intended to be "limiting cases" of a range of possibilities, in the sense of expressing "extreme weakness" of the two parties, respectively, and "dominance" of the other, in a sense implicitly defined by the behavior of the parties. For this range of prices and outputs the outcome is indeterminate. However, as will be seen presently, the weakness assumed in Cases (*a*) and (*b*), respectively, is "extreme" (or the "dominance" of the other party is complete) only if we exclude the possibility of imposing an all-or-none clause on the opponent. If all-or-none clauses are excluded, the complete lack of bargaining power expresses itself in the setting up of a market schedule (supply schedule or demand schedule, as the case may be). This is precisely what is implied in the discussion of Cases (*a*) and (*b*). Yet if all-or-none clauses *are* used by a party, then the other may do worse. He may be pushed down to the zero profit level. This is worse than having to set up a supply schedule or a demand schedule and letting the other party select a point along this schedule, because the point so selected is usually better than a zero profit point from one's own point of view.

Case (c) is characterized by the assumption that the two producers maximize their joint profits. This they will do by equating the marginal cost function of the seller with the marginal value product function of the raw material to the buyer (that is, by equating MC with MVP). The price — in contrast to the quantity — is "indeterminate": it may vary between the price shown by the average cost of the seller and that shown by the average value product function of the buyer, for the quantity equating MC with MVP. The sales contract will usually have to include an all-or-none provision. Otherwise, for any given price, the seller would prefer to deliver the quantity determined by the equality of MC with that price, while the buyer would prefer to buy the quantity

6. Let the aggregate value product of any quantity x of the raw material be $f(x)$. Then the average value product is $\dfrac{f(x)}{x}$, and $MVP = f'(x)$, $MMVP = f'(x) + xf''(x)$.

determined by the equality of MVP with the same price. Consequently, the maximization of joint profit *without* an all-or-none clause would be possible only at the price which equals MC and, at the same time, MVP. This actually is one of the prices lying between the average cost function of the seller and the average value product of the buyer, for the quantity that maximizes the joint profit. But it is merely one of these prices. For all other prices an all-or-none clause is required.

In Case (*c*) the outcome of the bargain is on Edgeworth's familiar *contract curve*. The contract curve is defined as the locus of bargains from which it is impossible to move towards another bargain so as to improve the position of one party without worsening that of the other. In other words, if we had represented the relevant data of the two firms by profit indifference curves,[7] instead of by cost and value product functions, then Case (*c*) would be marked by a definite stretch along the line connecting the points of tangency of the buyer's indifference map with that of the seller. The relevant stretch is limited by the zero profit level of the seller, on the one hand, and by that of the buyer, on the other. However, it is simpler to represent the problem with the aid of cost and value product functions.

The raw material output of Case (*c*) is the "competitive output" in the sense that the bilateral monopoly relationship *per se* does not distort the competitive pattern of allocation, provided it does not change the production functions. If in all *other* stages of the structure of production, selling and buying were purely competitive (which would imply that the monopsonistic buyer of the raw material was a competitive seller of *his* product), then the allocation of resources would be such as in pure competition. The same would be true if "Case (*c*) bilateral monopoly" existed in all stages in which products pass from one producer to another.

Figure I represents the theory here summarized. The arrows indicate the direction in which the value product functions, on the one hand, and the cost functions, on the other, shift, if both on the demand side and on the supply side new firms enter which produce

7. For the indifference curves in question, the raw material price would be measured along the ordinate, and the raw material quantity produced and sold, on the abscissa. The indifference curves would show given aggregate profits. The curves of the seller would be convex to the origin, those of the buyer concave.

close substitutes of the products of the two "original" firms and use similar factors of productions. The competition of these firms tends to shift the value product functions down and the cost functions up. The limiting situation would be characterized by an AC function that lies *above* the AVP function everywhere except at a
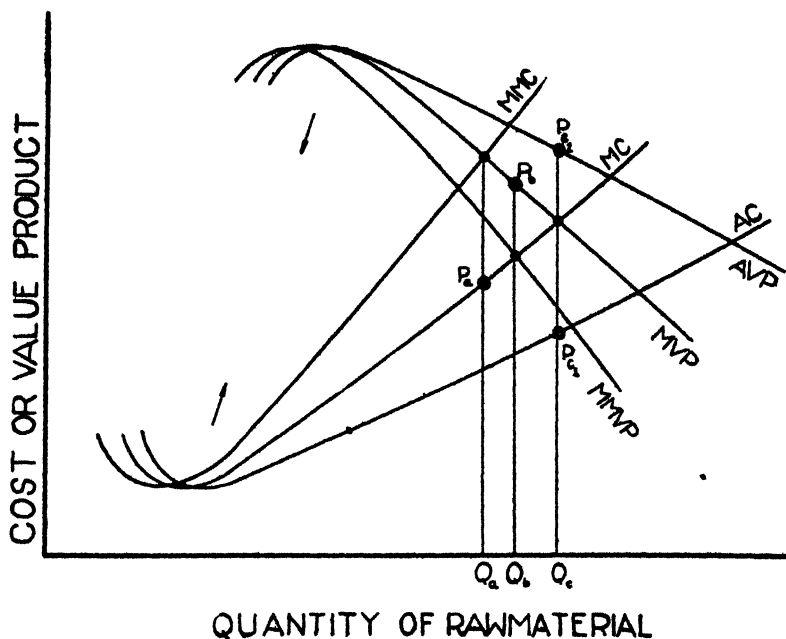
FIGURE 1

AC = Average Cost
MC = Marginal Cost
MMC = Marginal to the Marginal Cost
AVP = Average Value Product
MVP = Marginal Value Product
$MMVP$ = Marginal to the Marginal Value Product
Q_a = Quantity produced if buyer dominates
P_a = Price paid for raw material if buyer dominates
Q_b = Quantity produced if seller dominates
P_b = Price paid for raw material if seller dominates
Q_c = Quantity produced if joint profit is maximized
P_{c1} = Lower limit of price if joint profit is maximized
P_{c2} = Upper limit of price if joint profit is maximized

For precise definition of the functions, see footnote 4, p. 505 and footnote 6, p. 506.

single point of tangency. The MC and the MVP functions would then intersect precisely below the point of tangency. In other words, in this limiting situation, the equilibrium becomes determinate. Only Case (c) equilibrium is conceivable, because in all other cases at least one of the two firms suffers a loss; and, for the same reason, the price must be that corresponding to the tangency of AVP with AC. But this is merely a limiting situation of no practical significance.

We believe that the theory portrayed in Figure I calls for reconsideration in certain respects. However, it will be easier to develop our suggestions after having extended the theory to labor-management relations.

EXTENSION TO LABOR-MANAGEMENT RELATIONS: WAGE THEORY ON THE SPECIFIC EQUILIBRIUM LEVEL

In view of collective bargaining, labor-management relations typically contain an element of bilateral monopoly. They also contain important non-economic elements which cannot be approached from this angle, or can be brought in only indirectly. But the economic element, centering on wages and their equivalents, is significant enough to justify the attempt to apply the theory of bilateral monopoly to labor markets. The version of the theory summarized in the preceding pages is not *directly* applicable, because the behavior of the suppliers of labor services cannot be interpreted in terms of cost functions (in the usual sense of the term). It is useful to substitute an indifference map for the cost functions of Figure I. The shape of the indifference curves must reflect the assumptions which are made concerning the attitude of the labor representatives with respect to the advantages of high wages, on the one hand, and more employment (less unemployment), on the other. An attempt will be made to indicate briefly the main features of the model in which the problem of labor markets may be discussed.

An indifference map should be defined for the union, between wage rates (ordinate) and the quantity of employment (abscissa). The average and the marginal value product function of labor to the enterprise should be drawn into the same diagram (i.e. with the identical abscissa). The value product functions show the AVP and MVP of different amounts of labor input to the enterprise. These value product functions should be given an "inverse U-

shape."[8] At least three assumptions are worth distinguishing with respect to the indifference map.

Assumption I: the indifference curves are horizontal lines, and the indifference level for the union rises monotonically with rising wage rates. This means that the union is the better satisfied, the higher the wage rate, regardless of what the quantity of employment is. In this event, what is the upper limit of the bargaining range, i.e. what is the best the union could accomplish against a firm completely lacking "bargaining power"? Obviously, the union would aim at the wage rate (indifference line) which is tangential to the average value product function at the maximum point of the latter, and which intersects with the marginal value product function at the same point (i.e. at the maximum point of AVP). If the wage rate were raised higher, the firm would go out of business. The lower limit of the bargaining range, on the other hand, is determined by the wage rate (indifference line) below which it is impossible to push the union, either because at lower wage rates the workers in question prefer alternative employment opportunities,[9] or alternative types of organization, or because government regulations prevent the lowering of wages below this level. In reality, the wage rate will lie between the upper and the lower limits of the bargaining range, and it will be determined by the relative bargaining power of the union and the management, respectively. The amount of employment will be determined by the point of intersection of the wage line (indifference line) on which the two parties agree, with the marginal value product function.[1]

Assumption II: the indifference curves are horizontal lines, as on Assumption I, but the level of satisfaction does not rise monotonically with rising wage rates. Instead, it rises with wage rates up to a certain level (e.g. up to a wage level lying between W_{min} and W_{max} in the diagram in note 1, below), and then the level of satisfaction declines with a further rise in wage rages. This means that the union realizes vaguely the existence of some relationship between wage rates and the quantity of employment, and that for

8. The reader may wish to take a look at footnote 1, below, which will be explained presently.

9. The complications arising if, at a certain wage rate, *part* of the labor force prefers alternative opportunities and another part does not, will be discussed later.

1. The diagram on the next page illustrates the conditions discussed in the text.

this reason (and possibly also for other reasons) it prefers not to raise the wage rate beyond a certain but possibly vaguely defined limit. For the unemployment arising at "unusually high" wage rates (and possibly also for other maladjustments), the union would be made "responsible." But up to the wage level in question, it prefers higher to lower wages, regardless of the amount of employment. Then the upper limit of the bargaining range is set by the wage level just described.[2] The lower limit is the same as on Assumption I, above.[3] Employment will always be determined by the intersection of the wage line on which the two parties agree with MVP.

Assumption III: the indifference curves are concave from above, that is, they are shaped like the indifference curves applying
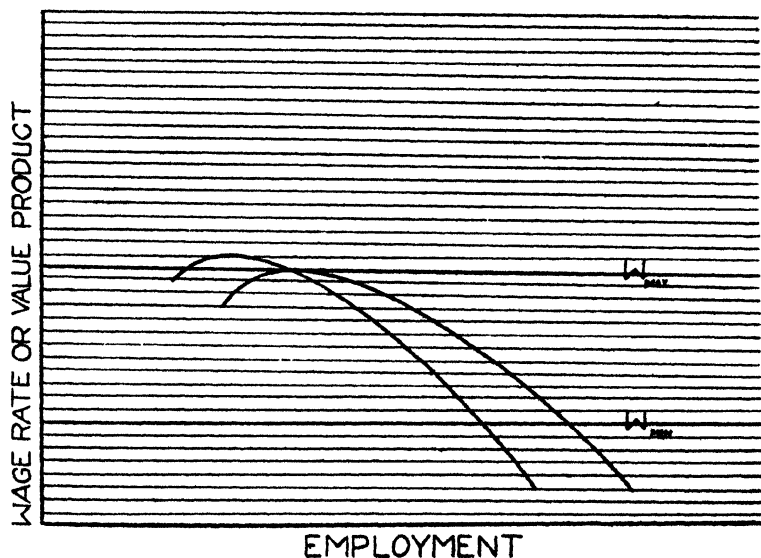
W_{max} = upper limit of bargaining range.

W_{min} = lower limit of bargaining range (determined by alternative opportunities or government regulation).

The two curves express the average and the marginal value product of labor.

2. Assuming that this wage level falls short of W_{max} in footnote 1, above, If it does not fall short of W_{max}, then W_{max} is the upper limit in this case, as on Assumption I.

3. Assumptions I and II are discussed also in the author's Monetary Policies and Full Employment. (Berkeley: University of California Press, 1946, Appendix to Chapter III.)

to two consumer goods. This means that the union realizes consistently that, throughout the range of possible outcomes, it is making a choice in which the amount of employment is involved as one of the variables, and that high wages and higher employment are sources of satisfaction for the union, standing to each other in the relationship of imperfect substitutes. Two subcases should be distinguished.

(Subcase I): it may be assumed that the labor contract relates *merely to the wage rate* (or, more precisely, that it does not extend to the amount of employment supplied). In this event, the highest conceivable bargaining power of the union would establish the wage rate corresponding to the tangency of the marginal value product function with the indifference map of the union.[4] In other words, that point of the MVP function would be realized which is the most favorable on the indifference map of the labor representatives. This statement should, however, be qualified by the proposition that if the point so located lies above the AVP curve, then even the maximum bargaining power of the union would establish a lower wage rate, namely, that corresponding to the intersection of MVP with AVP. The highest conceivable bargaining power of the firm would lead to the establishment of a wage rate corresponding to the intersection of the MVP function with the indifference level below which it is impossible to press the union. (The concept of this "lowest possible" indifference level was discussed on p. 510 in connection with Assumption I.) At any wage level employment will be determined by the equality of that wage rate with the marginal value product.

(Subcase II): if the *amount of employment* supplied is included in the labor contract — that is, if the contract contains an all-or-none clause — then the maximum bargaining power of the union could establish the wage rate corresponding to the tangency of the average (not marginal) value product function with the indifference map. The union would then enforce the amount of employment measured by the abscissa of this tangency.[5] The lower limit of the bargaining range would be set by the intersection of the *marginal curve to the "lowest possible" indifference line* with the marginal value product function. This intersection point determines the

    4. The diagram contained in footnote 5, below, will be helpful for visualizing these conditions.
    5. The distinction between the point of tangency of the indifference map with *MVP*, on the one hand, and with *AVP*, on the other, is illustrated by the diagram on the next page.

employment for the maximum bargaining power of the firm, and the corresponding point of the "lowest possible indifference curve" indicates the wage rate. In other words, the "lowest possible" indifference curve becomes the supply curve in these circumstances, and the firm chooses that point along the "supply curve" which is most profitable to itself. (This it cannot do in Subcase I, Assumption III, because in that subcase the labor contract established a horizontal supply function.)

Figure II shows the procedure by which the most profitable point along the given indifference curve is found in Subcase II.

The contract curve along which the bargain must lie, if an all-or-none clause is included in the contract, consists exclusively of points such as R, lying along each indifference curve and found with the same "marginal" method as was used to locate R on indifference curve LI in Figure II. The lowest-wage point on the contract curve is R in Figure II; the highest-wage point is that determined by the tangency of the indifference map with AVP (cf. the diagram in footnote 5 on p. 512). This latter point can also be found by the method used in Figure II, since it is the highest-profit point along the indifference curve which has a point of
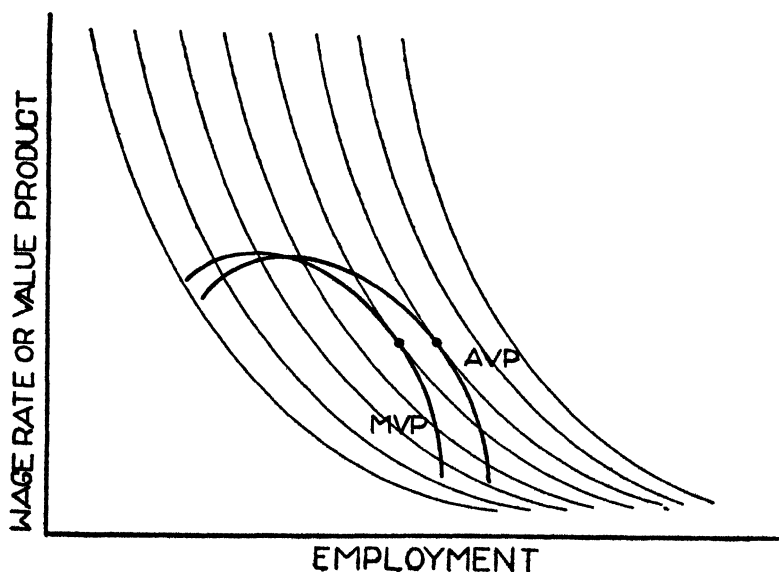
The "lowest possible indifference curve" is not marked in the diagram, nor is the marginal curve to this indifference curve drawn.

tangency with AVP. (All other points of this indifference curve correspond to negative profits.) Consequently, the intersection of the marginal curve to this indifference curve with MVP lies pre-
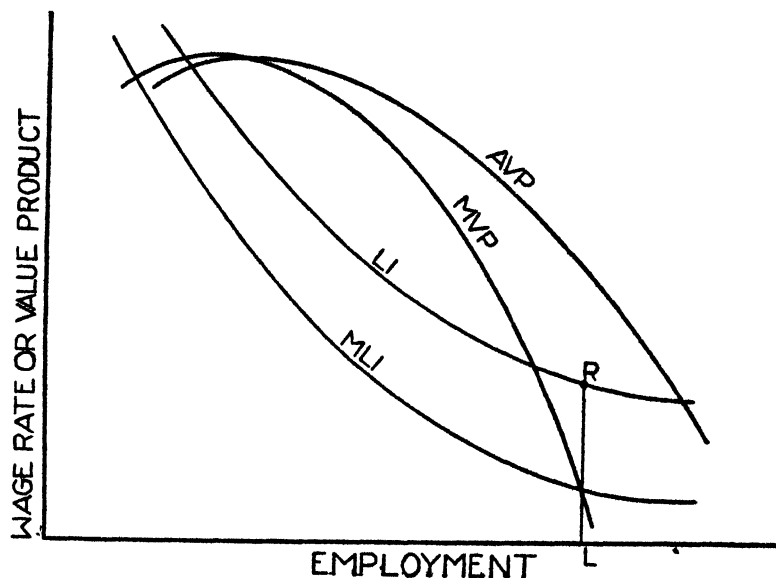
FIGURE II

AVP = Average Value Product Function
MVP = Marginal Value Product Function
  LI = Lowest possible Indifference Curve
MLI = Marginal curve to the lowest possible Indifference Curve
  LR = Corresponding wage rate

cisely below the tangency of the indifference curve itself with AVP. In general, this method is suitable for deriving the contract curve, because the method locates the highest-profit points along the subsequent indifference curves, and the contract curve consists of these highest-profit points.[6]

The conclusion may be drawn that wage contracts including all-or-none clauses offer certain advantages over "simple" wage-

6. Another suitable method would be that of constructing a profit-indifference map for the firm, consisting of indifference curves which show combinations of wage-rates and amounts of employment yielding a given aggregate profit. The contract curve consists of points of tangency of this profit-indifference map with the indifference map of the union.

rate contracts, *if* the union's indifference map is concave. For the all-or-none clause puts the parties on a point of the contract curve from which, by definition, it is impossible to move in any direction so as to improve the position of one party without worsening that of the other. The "simple"wage-rate contract, on the other hand, puts the parties on the intersection point of a horizontal wage line with MVP, and from such a point it *is* possible to move so as to improve the position of both (or of one without worsening that of the other). For example, in the absence of an all-or-none clause, the union could never agree to a wage rate as low as LR (Figure II), because for such a wage rate the firm would choose the employment determined by the intersection of MVP with the horizontal line through R, and this point lies below the lowest possible indifference level. Or, for the wage rate corresponding to the intersection point of LI with MVP, the firm would choose the abscissa of the inter-section point, itself, and this is no better for the union, but worse for the firm, than point R on Figure II. As Professor Leontief showed recently (with a somewhat different apparatus, but the same kind of reasoning),[7] the analysis proves that guaranteed annual wage contracts are superior in certain respects to ordinary wage contracts. The commonsense of the matter is that a lower wage rate coupled with higher employment may be preferable, notmerely to the firm but, in certain circumstances, also to the union. The advantage of the "all-or-none" disappears when the indifference map of the union consists of horizontal lines. *In this event the outcome of the bargain will lie along the contract curve, even if the contract relates merely to the wage rate.* The intersection of a horizontal wage line with MVP determines the same point as the intersection of the marginal curve to the wage line with MVP. In other words, with a horizontal indifference map, there is no difference between the method by which the amount of employment is determined for any given wage rate on Assumptions I and II (with no all-or-none clause) and the method by which the contract curve is derived for wage contracts including "all-or-none" propositions.

Even if the indifference map of the union is concave, the merits of an all-or-none clause cannot be judged on the basis of the foregoing considerations alone. Wage contracts are usually concluded for some considerable period of time in advance. Consequently,

---

7. Wassily Leontief, "The Pure Theory of the Guaranteed Annual Wage Contract," Journal of Political Economy, February, 1946, pp. 76–79.

employment guarantees, or wage guarantees, shift a substantial degree of uncertainty to the firms. This will ordinarily induce them to act as if their value product functions were lower than what corresponds to their best guess as to the profitability of their businesses at the moment when the contract is concluded. Such allowance for uncertainty[8] is, of course, a deflationary factor. If, on the other hand, they should prove to be too venturesome in estimating their value product function, they would become involved in difficulties, which also might adversely affect the economy as a whole. These disadvantages may frequently outweigh the advantages discussed in the preceding pages.

Some Complicating Factors

So far the analysis has assumed that, at all wage levels which might be set on the basis of the foregoing theory, enough individual workers are available to allow the firm to equate the wage rate to the marginal value product of labor. In a range of wage rates in which individual workers are available only in numbers insufficient to equate the wage rate with MVP, the traditional supply surve of labor becomes a significant concept. The supply curve may be upward sloping or backward rising, depending on whether the shortage decreases or increases with rising wage rates. If the supply curve is upward sloping, both the upper and the lower limit of the bargaining range *may* fall in the range of the shortage.[9] If the supply curve is backward rising, the lower limit *cannot* fall in the range of the shortage.[1] In case the upper limit falls in the range of shortage, the highest wage rate the union could obtain is that corresponding to the intersection of the supply curve with the average value product function. In case the lower limit falls in the range of shortage, the best the firm could accomplish is the monopsony wage rate, i.e. the wage rate indicated by the supply curve, for the amount of employment which equates the *marginal* labor cost function (that is, the marginal function corresponding to the supply curve) with the MVP function. For the range of shortage, it is

---

8. That is, for the kind of "risk" to which the probability calculus proper cannot be applied.

9. That is, the highest wage an unopposed union could obtain and the lowest wage an unopposed firm could establish might fall in the range in which not enough workers are available to equate the wage rate to *MVP*.

1. It follows from the theory of monopsony that with a backward rising supply curve the firm would find it profitable to lower the wage rate below the range of shortage.

true of any wage level on which the two parties might agree that the firm will hire all workers available at that wage level, provided the average value product of all available workers does not fall short of that wage rate. If the average value product of *all* available workers falls short of the wage rate, the same must also be true of a smaller number of workers.[2] Therefore, in this case the firm will go out of business. Needless to say, for the region of shortage Assumption III becomes irrelevant. The union will not be willing to make wage concessions for more employment.

Shortages give rise to complications for other reasons also. We were just now concerned with "shortages" in the sense that not all workers are available who could find employment on the terms of the agreement. But "shortages" may also exist in the sense that increasing amounts of employment can be provided, yet only at the cost of increasing disutility (overtime work). An attempt could be made to incorporate this consideration into the indifference maps, e.g. by drawing them with an upward slope for distinctly high levels of employment. While this might have certain advantages, it also has the disadvantage of requiring the introduction of further simplifying assumptions. The fact remains that an increase in labor input by means of increasing the number of laborers employed has different implications from an increase in input by means of overtime work. Including these two phenomena in the same two-dimensional model requires postulating that the one, and not the other, phenomenon takes place in definite regions of the map.

It seems preferable, therefore, to postulate that the model so far described relates to a working week of given length, and that, consequently, changes in employment mean changes in the number of workers employed. It becomes necessary, however, to add the following. If another wage-rate-employment indifference map is defined, with a given number of employed and changing hours,[3] then the union-indifference-curves in this *Supplementary Model* are always downward sloping as far as some "normal" working week, and from there on they are upward sloping. More employment is preferred for given workers up to the point where an "adequate" income is reached without "abnormally" long hours, and from there on more employment is accepted only at increasing wage

2. Because if the supply curve is to the *left* of MVP and *above* AVP, it must be above the *rising* range of AVP.

3. So that a change in employment (or in labor input) means a changing working week.

rates (overtime rates). This consideration determines the shape of any one union-indifference-curve in the Supplementary Model. As we move upward along the ordinate (wage-rate axis) of the Supplementary Model, the high indifference curves have their minimum points at lower employment levels than the low ones, because at high wage rates an "adequate" income is obtained with less work. Therefore, the length of the "normal" working week tends to decrease as wages rise, but not necessarily continuously. The shape of the union indifference maps in the Supplementary Model is indirectly determined by the attitudes of the employed labor force with respect to the advantages of leisure, on the one hand, and income, on the other. The Supplementary Model may be made to reflect one aspect of the phenomenon traditionally expressed in the "backward-rising supply curve," but in the present institutional setting it is impossible to present this aspect adequately in terms of individual supply curves.

Considering the circumstances expressed in the Supplementary Model, the firm may be said to draw from two alternative sources of labor supply. It draws labor from the "normal working hours-market" (Original Model) and from the "overtime market" just discussed (Supplementary Model). From the latter source, it will only draw if, due to indivisibilities, it cannot organize its activities so as to limit itself to the normal hours market. In this event, the residual will be drawn from the overtime market. A given level of satisfaction for the union corresponds to a higher wage rate on the overtime market than on the normal hours market. Consequently, given the relative bargaining power of the two parties, the overtime market is a more costly source of supply.

Returning to the Original Model, it should further be added that if the efficiency of labor itself depends on the height of the wage rate, different value product functions would have to be drawn for each potential wage rate. This would render the graphic representation of the problem difficult, but it is easy to indicate what kind of allowance would have to be made for this complicating factor, provided the relationship between wage rates and efficiency is known. In this case, the maximum bargaining power *of the firm* would not necessarily place the union on the lowest possible indifference level. It would place the union on a higher indifference level, provided the value product functions corresponding to this higher indifference level are higher by a sufficient margin

to make for greater profits. On Assumptions I and III, the solution for the maximum bargaining power *of the union* would still have to satisfy the tangency conditions developed in the preceding section of this article (cf. pp. 512, 513). But if different value product functions correspond to different wage rates, several wage rates may satisfy these tangency conditions (with the different value product functions corresponding to these wage rates). The highest of these wage rates will be chosen. On Assumption II it could merely be stated that a positive relationship between wage rates and efficiency is likely to increase the wage rate corresponding to the maximum bargaining power of the union. On the whole, the change required by the complicating factor here considered might be expressed as follows. The propositions of the preceding section of this article establish a relationship between wage rates, on the one hand, and hours of employment, on the other, *via* "the" (average and marginal) value product function. If a different (average and marginal) value product function corresponds to each wage rate, these propositions should be reinterpreted so as to relate to wage rates, on the one hand, and the specific value product functions corresponding to each wage rate, on the other. It should be remembered, however, that an upward shift of the productivity functions may require better wages in many industries at the same time, for it may require a change in the environment in which workers live. This phenomenon, which is in the nature of (potential) external economies, may be characteristic of backward areas. In these circumstances, wage rates may tend to be lower than would be in the interest of the employers (and, of course, also of the workers).

A further significant complicating factor, so far disregarded, arises from discontinuities in the value product functions. These result in ranges of indeterminateness within which wages may change without affecting the level of employment. The discontinuities in the value product functions may be caused by discontinuities in the production functions, or by discontinuities in the marginal revenue functions (cf., for example, the case of kinked demand curves). They are much more important in the short run than in the long run; and even in the short run, they are much less important for the economy as a whole, or for substantial segments of the economy, than for specific firms. Extension of the analysis to the economy as a whole gives rise, however, to problems of a

different character, which will be considered briefly in the next section of the article.

We have consistently disregarded the fact that the wage bargain relates to a variety of wage rates for different types of labor, rather than to a single wage rate, and also the fact that certain economic advantages may accrue to the workers, not in the form of wage payments, but in other forms. Furthermore, the foregoing presentation did not explicitly take into account the circumstance that one and the same wage agreement may relate to many firms. In principle nothing stands in the way of extending the analysis to these aspects of the problem, except that the exposition would become more involved. However, the problem of the non-economic objectives of firms and unions is a different matter. This problem could be brought in only indirectly, for instance, in a detailed discussion of the determinates of the two bargaining limits (cf. p. 510 and p. 512).

WAGE THEORY ON THE GENERAL EQUILIBRIUM LEVEL

Let us assume that wage rates change for the economy as a whole, with no *primary* change in the value product functions (e.g. in consequence of a change in relative bargaining power). In this event, the value product functions themselves will also change under the impact of changes in aggregate demand. This is one reason why a theory such as we presented in the preceding sections needs to be amended, if we wish to go outside the framework of the specific equilibrium approach. The other reason is connected with cross-elasticity relationships, with which we will not be concerned here.

An autonomous rise (fall) in money wage rates presumably tends to raise (lower) the value product functions because prices tend to rise (fall). Will the wage change be just offset by the shift in the value product functions, so that employment remains unchanged? Or will the wage change be less than offset by shifts in the value product functions, so that employment falls with rising wage rates and vice versa? Or will it be more than offset, so that employment rises and falls along with the money wage rate? An acceptable wage theory should contain an elaborate analysis of these possibilities, as well as a specific equilibrium discussion along the lines previously considered. We shall limit ourselves to making a brief suggestion. In developing it, we shall consider a wage

increase rather than a decrease, but all statements could be repeated for a fall in wages, provided the expressions "lowering" and "downward shift" are substituted for "rise" and "upward shift." *We shall assume that, after the rise in question, wage rates are expected to remain unchanged.*

In the event of an autonomous rise in money wage rates for the economy as a whole, certain forces are generated, among which we distinguish two. Force I tends to raise the value product functions *less* than what corresponds to the wage increase, and consequently it tends to reduce employment. In other words, the upward shift of the value product functions under the impact of Force I is less than is necessary to "offset" the wage rise, so far as the effect on employment is concerned. Force II produces a *further* upward shift in the value product functions, and consequently it tends to raise employment above the level to which Force I reduces it. In certain ranges the joint effect of the two forces may be sufficient to maintain aggregate employment at its initial level or even to increase it. In other ranges, this will not be the case.

Force I is generated by the circumstance that, in the economy as a whole, investment expenditure per unit of consumer demand is likely to fall when wage rates rise. Wage incomes (as a source of demand) possess a much more direct bearing on consumer demand than on the demand for investment goods. Part of the investment demand is not closely related to consumer demand, but, instead, is based on the expectation of such investment demand in the next period, which, in turn, will again be based on the expectation of investment demand in the following period, etc. For this kind of investment activity ("investment for further investment"), wages are costs, but not, in any direct sense, demand-influencing factors. Therefore, a general wage rise is likely to lower investment demand per unit of consumer demand. Aggregate consumers'-plus-investors' expenditure (i.e. aggregate income) will rise, but it will rise in a proportion smaller than that in which wage income and consumers' expenditure rise. Wage income will amount to a higher share of aggregate income. Consequently, real wage rates will tend to rise and employment will tend to fall.[4] This is how Force I operates.

---

4. A crude version of the Keynesian theory would maintain that real wages and employment remain unchanged. This also would be our conclusion *if we assumed that investment demand per unit of consumer demand remains unchanged.* In this case, Force I shifts the value product functions just in the proportion which is necessary to offset the wage rise, so far as the effect on

Force II is produced by the circumstance that Force I raises the average propensity to consume in consequence of the rise in real wage rates. Logically, a full employment economy is conceivable in which the average propensity to consume is very low, and in which most of the economic activity is investment for further investment, as described in the preceding paragraph.[5] (Aggregate consumption would, of course, still be rising with aggregate income, as long as the marginal propensity to consume is greater than zero.) Yet in such an economy the uncertainty attaching to profit expectation would be very high, and consequently a high level of employment would be unlikely to materialize. An increase in the average propensity to consume diminishes the uncertainty attaching to profit-expectations, because consumer demand is a more stable and "dependable" constituent of aggregate demand than is the demand for investment goods. Force II operates through this decrease in uncertainty, and it tends to increase aggregate output and employment. If the value product functions of labor are defined net of the costs of the coöperating factors, *including the costs of bearing uncertainty*, then the effect of Force II may be expressed by saying that a further upward shift is produced in the net value product functions. If we allow for the decrease in uncertainty in some other fashion, then a different terminology is required at this point, but the outcome of the analysis remains the same. A decrease in uncertainty tends to increase the level of employment.

Somewhat crudely expressed, the analysis suggests that the aggregate effect of a general wage rise depends on whether the cost-raising effect generated by Force I is stronger or weaker than the propensity-to-consume-increasing effect (or uncertainty-decreasing effect) generated by Force II. The joint effect of Forces I and II is much more likely to be sufficient to maintain, or even to increase, employment in a range in which the average propensity to consume is low and in which profit margins are high, than in a range in which the average propensity to consume is high and in which profit margins are low. Force I, alone, results in decreased employment, and Force II is likely to be weak in circumstances

employment is concerned. Moreover, if this happens, Force II is not generated.

5. For detailed discussion, see the present author's Monetary Policies and Full Employment, Chapter II and passim.

characterized by a high average propensity to consume and by low profit margins.

RECONSIDERATION OF THE THEORY OF BILATERAL MONOPOLY

When we reviewed the "present state" of the theory of bilateral monopoly, we distinguished Cases: (a) where the seller is "weak" and the buyer dominates; (b) where the buyer is "weak" and the seller dominates; and (c) where the seller and buyer maximize joint profits. We also pointed out that Cases (a) and (b) should be regarded as limiting cases for a whole range of potential equilibria. In the present section we shall argue that this theory calls for reconsideration. Before discussing this question, however, we need to inquire how the statements contained in the sections concerned with the labor market fit into the general framework with which we started out in the first section.

With the exception of one situation envisaged in the section on wages, the entire wage discussion is essentially a discussion of Case (c) equilibria. The exception is the situation considered under Assumption III, Subcase I, that is, the situation in which the union indifference map is concave, but in which the contract nevertheless does not extend to the amount of employment.[6] Aside from this possibility, all points that could express the outcome of wage bargains lie along "contract curves." In other words, all possible contracts place the parties on contract curves if the problem belongs under Assumption I or II, or if it belongs under Assumption III, Subcase II, as discussed in the section on wage theory (specific equilibrium level).[7] In all these situations a point is realized from which it is impossible to move towards greater satisfaction for one of the two parties without diminishing at the same time the level of satisfaction of the other party. This condition is satisfied by the intersection points of the horizontal union indifference lines with the MVP curve, that is, by the points which express the possible outcomes of the bargain on Assumptions I and II. The same condition obviously is satisfied by the locus of possible outcomes on Assumption III, Subcase II, because bilateral monopoly contracts with an all-or-none clause will always satisfy this condition (except for "mistakes" with which we are not concerned here). Consequently,

6. Cf. p. 512.
7. Cf. pp. 512–513.

in all these cases, we are faced with Case (c) equilibria, as discussed in the section on the "present state of the theory," even though no all-or-none clause is postulated on Assumptions I or II. The characteristic property of Case (c) is that it is impossible to improve the position of one of the two parties without worsening the position of the other party.

This brings us to our main point of criticism. Cases (a) and (b), for product markets, and Assumption III, Subcase I, for the labor market, do not satisfy the crucial condition in question. They establish a set of outcomes from which it *is* possible to move in some direction so as to improve the position of both parties at the same time. In other words, these outcomes do not lie along "contract curves," and hence it would always be in the interest of both parties to move from such an outcome to some point along the contract curve which can be defined for the problem under consideration. Is it legitimate to assume that the parties will not do so? Are Cases (a) and (b)[8] for product markets, and Assumption III, Subcase I, for the labor market, "legitimate" cases?

The answer depends on whether it is reasonable to assume that institutional obstacles prevent the parties in question from moving to the contract curve (that is, from establishing Case (c) instead of (a) and (b) for product markets; or from establishing Assumption III, Subcase II, instead of Assumption III, Subcase I, for the labor market). We submit that, for the labor market, it is reasonable to "salvage" Assumption III, Subcase I, with reference to institutional factors, but that, for product markets, it is unrealistic to try to salvage Cases (a) or (b), or the range for which these are the limiting cases. On the labor market, institutional circumstances typically do prevent the realization of Subcase II, instead of Subcase I, provided the problem belongs under Assumption III at all (i.e. provided union indifference maps are concave.)[9] But it is difficult to conceive of institutional obstacles that would prevent the realization of Case (c), instead of the other cases, on product markets.

The institutional obstacle, if it exists at all, must be an obstacle preventing the inclusion of an all-or-none clause in the contract.

8. Or the cases belonging in the range for which these are the limiting cases.

9. If the problem belongs under Assumptions I or II, the difficulty here discussed does not arise. The outcome will be on the contract curve anyway.

Such an obstacle usually *does* stand in the way of moving to the contract curve, if union indifference maps are concave. The period for which wage contracts are concluded is long enough to make it highly desirable for the management to avoid committing itself to a definite amount of employment. From the point of view of the management, the significance of uncertainty considerations outweighs the advantage of moving to the contract curve by means of an all-or-none clause. Without the all-or-none clause, the locus of possible agreements is different from those shown by the contract curve: we are faced with Assumption III, Subcase I, instead of Assumption III, Subcase II. On Assumptions I and II, the outcome is on the contract curve, even without an all-or-none clause, and we are inclined to believe that Assumption II is more realistic than either Assumption I or Assumption III. Yet Assumption III cannot be excluded, and in the framework of Assumption III it is impossible to exclude Subcase I. While in the "instantaneous" sense both parties would gain if they established Subcase II, i.e. if they moved to the contract curve by means of an all-or-none clause, uncertainty considerations typically prevent the inclusion of such a clause in wage contracts.

Not so for product markets. Both parties gain at the same time if they move from Case (*a*) or (*b*)[1] to Case (*c*) by means of an all-or-none clause, *and ordinarily there is nothing to prevent them from so doing*. For an all-or-none clause is implicit in the passing of a product from one firm to another in all cases except those in which a price is "quoted" regardless of the quantity that will be taken at that price. All that is required to establish the equivalent of an all-or-none clause is *not* to quote prices for indefinite quantities. In the business relations existing between a monopsonist and a monopolist for a product, nothing ordinarily stands in the way of quoting prices for definite quantities. Whenever price offers relate to definite quantities, an all-or-none clause is implicit in these offers and in the contracts based on them. Joint profits are maximized, and it may be added that they are maximized by a procedure that could scarcely be termed collusion.

We may conclude, therefore, that for product markets Cases (*a*) and (*b*) should be disregarded, and that bilateral monopoly results in a definite ("determinate") output, such as is estab-

1. Or from the range of outcomes limited by these two cases.

lished in Case (*c*). This output is the "competitive" output, in the sense of being determined by the equality of the marginal cost of the seller with the marginal value product to the buyer. If the buyer-monopolist were a competitive seller,[2] then the quantity of the finished output would be the "ideal" quantity. In this case the marginal value product is at the same time the value of the physical marginal product. If simple monopoly, monopsony, oligopoly or oligopsony exists in the structure of production prior to the stage in which bilateral monopoly occurs, or following this stage, then the finished output will be different from the "ideal" (competitive) output, but the bilateral monopoly itself — in whatever stage it may occur — will not contribute to this divergence, provided it occurs on product markets.

On product markets, bilateral monopoly output is determinate; price, however, is "indeterminate," in the fashion indicated for Case (*c*). This statement presupposes that the parties concerned are well informed and "rational," in the same sense as is implied in all "marginal revenue-marginal cost" propositions of value theory. These implications never are entirely realistic. Sometimes they may be quite unrealistic. But the qualifications arising from this circumstance, important as they are, have nothing to do with the problem considered in the present article. Cases (*a*) and (*b*) do not differ from Case (*c*) in that they take into account the general qualifications to which equilibrium theories are subject; they are "established" by the same *type* of theorizing as Case (*c*). The point is that the specific assumptions on which Cases (*a*) and (*b*) are derived do not possess internal consistency.

On the labor market, output and price are "indeterminate" (i.e. they both depend on bargaining power), as was established separately for Assumption I, Assumption II, and for Assumption III, Subcases I and II (pp. 510–513). Deviations from the "competitive" input must, however, be said to exist even on Assumptions I and II and on Assumption III, Subcase II, although in these cases the outcome of the wage bargain is on the contract curve. On product markets, if the bargain establishes a point on the contract curve, as in Case (*c*), the output is determinate, and it may be regarded as identical with the "competitive output" in the sense

2. Or if his product also were sold in the framework of a further *bilateral* monopoly relationship, in which he is the seller-monopolist and in which another producer is the monopsonist.

just mentioned (p. 526). But this is not true of the labor market. The difference between the conclusion for product markets, on the one hand, and the labor market, on the other, is a result of the following considerations.

When, on a product market, an output, determined by the equality of marginal cost and marginal value product, is regarded as the "competitive output," a generally accepted "convention" is adopted.[3] The ultimate implications of this convention are subject to a definite qualification, which is well-known and which has frequently been discussed (and, therefore, should not be misleading). The qualification is that the production functions are not altered by the (imaginary) change from pure competition to monopoly. It is generally understood that this is the "ceteris paribus" clause implied in the terminology by which a certain monopoly output on product markets is considered "equal to the competitive output." While the realism of such a ceteris paribus clause is questionable indeed, it is not obvious in what direction the actual deviations go, and it would be unnecessarily pedantic to point out consistently that the monopoly output in question is equal to the "competitive output" *only if unchanging production functions are assumed.* This is generally understood. Yet the analogous ceteris paribus clause for the labor market would be an absurd clause. Any terminology implying such a clause would be absurd terminology, rather than merely "conventional" terminology. The clause, as applied to the labor market, would have to maintain that the scale of preferences, as between wages and employment, from which the supply-behavior of the labor force as a whole is derived in competitive conditions, is the same as the scale of preferences from which unions derive their supply-behavior. *This* would be the analogy to the ceteris paribus clause, which on product markets applies to production functions rather than to psychological indifference maps. Less precisely, but more briefly expressed, the ceteris paribus clause, as applied to labor supply, would have to maintain that unions represent the interests of the unemployed and of the outsiders *just as well* as they represent the interests of their own employed members.

Such a clause should not be implied. If "repercussions" are

3. Equality of marginal cost with marginal value product also means equality of marginal cost with value of the marginal physical product, *if* the buyer, in turn, sells *his* product in purely competitive conditions. If he does not do so, the deviation from the competitive allocation is a consequence of *this* circumstance, which is distinct from those here considered.

included in the analysis, unions may frequently turn out to serve indirectly the interest of "the economy as a whole," not merely those of their employed members, because they may prevent the extreme instability developing from a very low average propensity to consume. But this does not mean that the scale of preferences on which they act represents those of all individual workers alike. For the union the "level of satisfaction" derived from a one-dollar wage rate plus 90 per cent employment is clearly different from what this wage plus this degree of employment means to the employed 90 per cent, on the one hand, and to the unemployed 10 per cent, on the other. The preferences "of the union" are in the nature of weighted averages, not in the nature of marginal preferences for the labor force as a whole.

The argument may be summarized as follows. On product markets, bilateral monopoly tends to establish a determinate output which equals "competitive output" (for the stage of production in which it occurs), but the price stays "indeterminate," as in Case (*c*). The outcome of the bargain tends to lie on the contract curve. On the labor market, the bargain also tends to lie on the contract curve — except on Assumption III, Subcase I — but this does not mean that the "competitive input" is established. It merely means that the outcome is optimal *for an aggregate consisting of the firm and the union*. Input as well as price (i.e. wage rate) are "indeterminate." They both depend on bargaining power.

The Contract Curve and the Concept of the "Solution"

It seems worth while to restate the argument briefly in terms of a concept introduced recently by Professors von Neumann and Morgenstern in their *Theory of Games*. We mean the concept of the "solution."[4] This concept appears to be of wide applicability. Consequently, it is useful to state a specific argument, such as ours, in these broader terms. This will have to be done sketchily.

Call any distribution of a total gain among the participants of a "game" an *imputation*. Adopt the following terminology: an

4. Cf. John von Neumann and Oskar Morgenstern, the Theory of Games and Economic Behavior, Princeton, 1944, pp. 37ff.; also the following review articles: Leonid Hurwicz, "The Theory of Economic Behavior," American Economic Review, December, 1945, and J. Marschak, "Neumann's and Morgenstern's New Approach to Static Economics," Journal of Political Economy, April, 1946. These two review articles were reprinted as Cowles Commission Papers, New Series, No. 13.

imputation (i_1) dominates another imputation (i_2), if (i_1) is more favorable than (i_2) to a number of participants which is sufficient to enforce (i_1). Finally, define a *solution* as a set of imputations satisfying the following conditions: (1) no imputation belonging in the set in question dominates any other imputation belonging in the same set; (2) any imputation *not* belonging in the set in question is dominated by some element belonging in the set. A "game" may have more than one solution, in which case we say that the various solutions correspond to different *standards of behavior*. No solution is "rationally" superior to another solution, just as, in the framework of any one solution, no imputation is "rationally" superior to the other.

The argument of the present article may now be stated in a few sentences as follows. Both for the "product-market" problem and the "labor-market" problems previously discussed, a unique contract curve exists in each case. This is tantamount to saying that a unique solution exists in each case, consisting of a definite stretch along the contract curve. This is a stretch limited by the postulate that neither of the two participants must be pushed below some definite indifference level (which essentially always is a long-run zero-profit level). Such a stretch along the contract curve forms a set of imputations which satisfies the requirements of a "solution," as defined in the preceding paragraph.

Case (c) for product markets, and Assumptions I and II, as well as Assumption III, Subcase II, for the labor market, are consistent with the criteria here considered. These are "contract curve" cases. However, Cases (a) and (b) for product markets,[1] and Assumption III, Subcase I, for labor markets, are inconsistent with these criteria. Our conclusion was that Cases (a) and (b) (for product markets) should actually be disregarded on these grounds.[5] Assumption III, Subcase I (for the labor market) should not, however, be ruled out.

Assumption III, Subcase I, may be established (in spite of the apparent "superiority" of Subcase II), as a consequence of certain institutional factors which stand in the way of the inclusion of all-or-none clauses in labor contracts. This is not analogous to the conditions existing on product markets, because on these markets the contracts typically *do* include an implicit all-or-none clause (*i.e.* price offers typically relate to definite quantities).

5. As well as the range of cases for which these are the limiting cases.

The "exceptional case" we have discussed (that is, Assumption III, Subcase I), for the labor market is, however, more apparently than really exceptional. The "exceptional" character of the conclusion is a consequence of the fact that the model in which the problem was analyzed is not fully adequate for the analysis. The model from which the specific industry equilibrium was derived disregards the institutional factors in consequence of which labor contracts are always concluded for considerable periods of time; and it also disregards the element of *uncertainty*,[6] due to which all-or-none clauses are not usually included in these contracts. Consequently if, before conclusions are reached, these institutional factors and elements of uncertainty are taken into account separately, the outcome is different from what it would be in terms of the formal model. The contract curve applies to the formal model, in which actually Subcase II (rather than Subcase I) would have to be expected, provided Assumption III is valid. The elements disregarded in the model usually deprive the contract curve of its significance and give "Subcase I" results, which are non-contract-curve results and, therefore, appear to be strange in terms of the formal model.

Certain institutional factors and the element of uncertainty are disregarded also in the specific industry models relating to product markets. But here these circumstances do not tend to produce "Case (a)" or "Case (b)" results, instead of the "Case (c)" solution. They merely make it necessary to emphasize the general qualifications to which all equilibrium theories are subject.

THE RANGE OF INDETERMINATENESS

We have seen that on product markets price (but not output) is "indeterminate" in a certain range, and that on the labor market both the wage rate and the input remain "indeterminate" between two limits. Now the statement that certain prices, wage rates, or inputs are indeterminate within given bargaining ranges merely means that the type of theory which has been applied is not suitable for explaining the processes that actually produce determinate results within the ranges in question.

A simple proposition will be submitted from which certain conclusions may be drawn directly. Assume that each party starts

6. That is, of the kind of risk that is uninsurable because the probability calculus in the sense proper cannot be applied to the universe in question.

from his own end of the bargaining range.[7] The willingness of each party to move towards the other[8] depends on his appraisal of the probability that the other will *not* move. To say that a party is willing to move beyond a certain point is merely another way of saying that he believes that the other party will not move as far as the point in question (from the other end). The appraisal of the opponent's willingness to move results partly from an appraisal of the "real" factors affecting his position, and partly from an appraisal of his psychological characteristics. The following are some of the conclusions suggested.

1. A permanent stalemate develops only if *each* party overestimates the willingness of the other to move. If *either* of the two parties did not overestimate the other party's willingness to move, he would move himself. Moreover, for a permanent stalemate, it is necessary that *each* party should continue overestimating the other's willingness to move, in spite of the fact that, for a period of increasing duration, he already has had the experience that the other party's willingness to move has been smaller than he had expected. This is not a logically impossible set of assumptions, but it is a very unlikely set. Consequently, in the usual kind of business relations, a permanent stalemate is very unlikely.

2. However, a permanent stalemate is less unlikely in cases in which something in the nature of a *deus ex machina* is expected from the stalemate. This *deus ex machina* usually assumes the form of a non-economic process, which may be started by the existence of the stalemate itself, and which, once started, is likely to change the basic relationship in which the two parties stand to one another. If the stalemate is expected to touch off such a process, the likelihood of a stalemate is increased. In these circumstances a stalemate might develop, and the non-economic process may be touched off, if only *one* party overestimates the likelihood that the process in question will change the basic relations in his favor. *It is not necessary that both parties should be too optimistic.* Consequently, a permanent stalemate is more likely (less unlikely) to develop in negotiations that have significant political implications than in the usual kind of business relations. (Yet, even if this were not the case, the "bilateral monopoly type of

7. That is to say, from the point of the range which is most favorable to him.

8. That is, to make concessions.

relationship" would still be inherently dangerous, whenever it applies to vitally important areas of human contact. A permanent stalemate is "unlikely," but it is bad enough to have major wars, or even general strikes, "once in a while.")

3. If each party underestimates, rather than overestimates, the other party's willingness to move towards him, then an agreement will surely be reached. Both parties will actually move. Even if only one party underestimates the other party's willingness to move, an agreement will be reached, and it will be favorable to the party whose willingness to move is underestimated.

4. If one of the two parties estimates the other party's willingness to move correctly, an agreement will be reached, regardless of whether the estimate of the other party also is correct.

5. A detailed description of the process becomes complicated, mainly in consequence of the fact that the behavior of each party during the negotiations is usually intended to change the other party's willingness to move. But this does not invalidate the foregoing propositions.

WILLIAM FELLNER.

UNIVERSITY OF CALIFORNIA

COLLECTIVE BARGAINING BY AIR LINE PILOTS

SUMMARY

Introduction, 533. — Growth of ALPA, 533. — Government of ALPA: membership, 538; councils, 539; the convention, 539; officers, 541; co-pilots, 542; affiliated groups, 544. — Salaries and Earnings: the Pre-NRA period, 545; Decision 83, 546; co-pilots' pay, 551; international service, 551; the Stratoliner arbitrations, 553. — The Four-Engine Dispute, 554. — Hours of work and mileage limitations, 562. — Rules and working conditions, 564. — Theoretical bases of high pilot pay: qualifications, 569; earning life, 571; hazards, 572; technological advance, 573; bargaining power, 574. — Conclusion, 575.

Aided by generous mail subsidies, which were inaugurated in 1926 when private contractors took over mail flying from the Post Office, and which have been continued by subsequent legislation including the Civil Aeronautics Act of 1938, the basic air transport law today, scheduled air transportation has developed into an industry employing approximately 75,000 persons, of whom about 7,000 are pilots or co-pilots.[1] In addition to the main regularly scheduled carriers, there are several other groups in the industry: "feeder" lines, which operate under Civil Aeronautics Board certificates, carrying passengers, mail, and cargoes from a fixed base on a regular scheduled route to various surrounding points; intrastate lines, which operate under state laws if such exist in their territory; and unscheduled or chartered carriers, which have been exempted from general CAB regulation till recently.

THE GROWTH OF ALPA

Organization among air-line pilots took form early in the industry's history; but although one "strike" occurred in 1919 — actually it was a protest of pilots against the insistence of Post Office officials that the mail must go through on schedule regardless of weather — the early organizations were largely fraternal rather than collective bargaining institutions.[2] The depression commenc-

---

1. For a discussion of this legislation, see C. E. Puffer, Air Transportation, Philadelphia, 1942. It should be noted that air transport is far from unique in having received government aid. Virtually all forms of transportation have so benefited in a variety of ways.

2. These early fraternal groups were the Air Mail Pilots of America, organized in 1920, when the pilots worked directly for the Post Office. Under private management in 1926, the pilots formed the National Pilots' Association.

ing in 1929, plus the introduction during that period of bigger and faster planes accompanied by a changeover from a system of base pay plus mileage to either a flat monthly salary or base pay plus an hourly rate without the mileage factor, which reduced earnings, was the immediate cause of the organization of the Air Line Pilots Association in 1930 by a group of veteran pilots under the leadership of David L. Behncke. By 1932, the ALPA included more than three-fourths of the scheduled line pilots in the United States.[3] It affiliated with the American Federation of Labor in 1931.

Early in 1932, the management of Cord's Century Air Lines announced a further reduction in pilot pay, proposing to use the savings to carry air mail at a lower figure than was then in effect. The pilots refused to fly at the rate, and a strike occurred. This marked the first in a long line of successful moves by the ALPA to invoke the aid of the Federal Government in a collective bargaining dispute. Hearings on the Century proposal before the Committee on Post Offices and Post Roads of the House of Representatives in March, 1932, were soon turned into a forum for the consideration of the strike. Since no federal agency wished to take the responsibility for assisting the air line, in view of the danger of public reaction from any accidents which might occur with strike-breaking pilots, the net effect of the controversy was that the Cord Corporation was forced to liquidate its air line subsidiaries. Meanwhile, the House Committee adopted a resolution informing the Postmaster General of its desire that collective bargaining privileges be guaranteed to the employees of all air lines to whom air mail contracts were let in the future.

The ALPA then had bills introduced into Congress which would have placed the air transport industry under the Railway Labor Act. The bills were opposed by the industry and the Post Office Department, however, and did not reach a vote.

The passage of the National Recovery Act in 1933 kept the ALPA-carrier relations before government forums. To circumvent an industry-proposed minimum of $250 per month, the ALPA secured exemption of pilots' wages and hours from the code on the

In addition, a West Coast group, the Professional Pilots' Association, existed during the late 'twenties. (Air Line Pilot, April 5, 1932.)

3. For the early history of ALPA, see Federal Coördinator of Transportation, Hours, Wages and Working Conditions in Scheduled Air Transportation, 1936, pp. 98–115; and Oscar Leiding, "Pilot Union President," Air Transport, November, December, 1943, and January, 1944.

ground that they were professional employees. Thereupon, five air carriers,[4] who had held new salary scales in abeyance pending the adoption of a NRA code, announced their intention of making them effective on October 1, 1933, coincidentally with the introduction of higher-speed planes. Since these new scales marked the complete abandonment of ALPA favored mileage pay in favor of a base plus hourly rate, a strike was called for midnight September 30, 1933. It was averted when the NRA National Labor Board assumed jurisdiction.

While the case was being adjudicated, the Post Office cancelled all air mail contracts. The Board thereupon prepared its decision, which it held in abeyance until May 10, 1934, when Congress appeared to be ready to adopt a new air mail policy. It then issued its famous Decision No. 83, which provided for pilots' salaries composed of base pay, plus mileage pay, plus hourly pay; bonuses for night flying; the maintenance of co-pilot salaries and terrain pay then in effect; and a flight limitation of eighty-five hours per month.

Because of the cancellation of air mail contracts, Decision 83 was of no immediate effect. Congress, however, soon made it applicable to all domestic air lines, rather than just the five carriers involved in the wage dispute, when at the request of the ALPA it included the following provision in the Air Mail Act of 1934:[5]

It shall be a condition upon the awarding or extending the holding of any air mail contract that the rate of compensation for all pilots, mechanics and laborers employed by the holder of such contract, shall conform to decisions of the National Labor Board. This section shall not be considered as restricting the right of collective bargaining on the part of any employee.

A Government air mail contract is so important for the successful operation of an air line that most domestic air carriers quickly adopted the Decision 83 scale, but in a few cases revocation of air mail contracts by the Post Office was threatened before compliance was achieved.

The original decision of the National Labor Board would have expired within one year, but on August 14, 1935, Congress amended the Air Mail Act of 1934 to require contracting pay scales for "all pilots and other employees" to conform to "decisions heretofore or hereafter made by the National Labor Board or its successor in

4. United, Eastern, Transcontinental and Western, American, and Western.

5. Public Law, No. 308, 73d Congress, Sec. 13.

authority, notwithstanding any limitations as to the period of its effectiveness included in any decision heretofore rendered." [6] A similar provision was contained in the Civil Aeronautics Act of 1938, which is the basic law governing air transport today.[7]

The ALPA's continued desire to be included under the Railway Labor Act resulted in legislation passed in 1936 with the full support of organized labor and the Roosevelt Administration. The protection given by the Railway Labor Act has been increased by a provision in the Civil Aeronautics Act of 1938 to the effect that compliance with the provisions of the Railway Labor Act shall be a condition for granting of a certificate of convenience and necessity by the Civil Aeronautics Board, which all scheduled air lines must have in order to operate.

Having achieved the highest minimum wage and lowest maximum hours legislation in the United States, if not in the world, the ALPA, slowed by its own centralized contract-making machinery,[8] took its time in codifying its bargaining arrangements with the various air carriers. The latter put no obstacle in the path of writing contracts, not one having ever compelled the ALPA to prove its right to represent pilots and co-pilots by invoking the election machinery of the National (Railway) Mediation Board. The first general ALPA agreement was entered into in 1939 with American Airlines, and subsequent agreements have been made with all regular domestic certificated carriers, except All-American Aviation, Inc., which operates a mail and express pick-up service in Ohio and West Virginia, and a number of newly certificated feeder lines.[9] The ALPA also has agreements with all American-owned certificated carriers in international service, including Pan American-Grace Airways, Inc., the leading Civil Aeronautics Board certificated American air carrier in South America. It also claims to represent a majority of the pilots of Alaska Airways, which operates in that territory, but it has no agreements with

6. Public Law, No. 270, 74th Congress, Sec. 11.
7. Public Law, No. 706, 75th Congress, Sec. 401 (L).
8. Its by-laws centralize agreement-making authority in the president's hands. See below, pp. 541–542.
9. On January 1, 1947, ALPA had agreements with the following carriers: American, American Overseas, Braniff, Chicago and Southern, Colonial, Continental, Delta, Eastern, Inland, Mid-Continent, National Northeast, Northwest, Pan American, Pan American-Grace, Pennsylvania-Central, Transcontinental and Western, United, and Western. A contract with Pioneer is being negotiated.

Hawaiian Airways which operates in that island group; with Caribbean Atlantic, a small line operating in Puerto Rico and the Virgin Islands; with TACA (Transportes Aereos Centro Americanos, Inc.), the largest Central American air carrier, which is not certificated by the CAB but by the countries in which it flies, or with any intrastate lines. Today ALPA membership is approximately five thousand.

Most foreign air carriers operating into the United States have their own pilot unions. The ALPA has agreements with two of these, the British Air Line Pilots Association and the Canadian Air Line Pilots Association, which provide for close coöperation, and it has approached the Dutch, Scandinavian and French pilots for similar arrangements.[1] In South America, it has on occasion aided groups of pilots operating for non-United States lines, many of whom use United States pilots.[2]

ALPA has no agreements with the many charter and non-certificated companies. In this field, it is being challenged by the Military Pilots' Association, which was organized in 1945 by former World War II armed service flyers, most of whom either operate non-scheduled passenger or cargo service, or are employed by carriers in this field. The MPA has also assisted former army and navy flyers to secure jobs, and has attempted to gain seniority credits for them on scheduled air lines for time flown in the armed services. The MPA has attributed its failure to achieve any success in accomplishing the latter task to the ALPA's refusal to coöperate. Perhaps for this reason, MPA officials made thinly veiled offers to operate TWA planes during the 1946 strike.[3] Besides these activities, the military pilots group also acts as a pressure body for the non-certificated carriers. Its claim of fifty thousand members, of whom only a small fraction are active, is probably greatly exaggerated.

The ALPA has recently shown considerable interest in other classes of air-line employees. An assessment was levied on the membership in 1945 to launch the Air Line Stewardesses' Association, which has won bargaining rights on several carriers, and is now functioning as an ALPA affiliate. ALPA leaders have also

1. The text of the agreements between ALPA and the Canadian and British groups is found in Air Line Pilot, April, 1943.
2. See Air Line Pilot, December, 1942; October, 1943.
3. American Aviation Daily, October 26, 1946, p. 271. For ALPA retaliation, see below, p. 544, note 5.

attempted to induce unions composed of navigators, flight engineers and radio officers to affiliate with it and to form an affiliated mechanics union. It has been unsuccessful because it has offered proposed affiliates only auxiliary status with no voice in ALPA policy; and because many non-pilot employees regard pilots as "aristocrats" who have little interest or concern for other groups.[4]

THE GOVERNMENT OF ALPA

The government of the ALPA is of special importance in explaining various policies of the organization. Hence it will be examined in detail.[5]

Membership in the ALPA is open to "any person . . . who serves as a first pilot, reserve pilot, or co-pilot."[6] However, an applicant must serve one year of active membership in a piloting capacity before his application becomes final.[7] Members are divided into four classes: active, inactive (i.e. those who have retired from duty as pilots), apprentice (i.e. those "who possess all requirements of membership with the exception of having served an aggregate of less than one year in an air line piloting capacity"), and honorary.

Initiation fees until recently were $100 for first pilot, $50 for reserve pilot, and $25 for co-pilot; annual dues, $100 for first pilot, $60 for reserve pilot, and $28 for co-pilot. The 1947 Convention modified these somewhat by making them dependent on individual earnings so that senior men now pay more, junior men less. ALPA laws also provide that "assessments may be levied on all members, pro-rated in the same ratio as the annual dues, to take care of

4. The TWA strike of pilots in 1946 caused considerable resentment among the lower paid groups who lost three weeks' pay.

5. Quotations from the ALPA laws are from the 1944 edition, by permission of David L. Behncke, president and copyright owner.

6. Reserve pilots are qualified to act as first pilots, but do not have a permanent first pilot run. Hence they fly part of the time as first pilots, and part as co-pilots. Reserve pilots are classified for ALPA voting purposes as first pilots, if they fly in excess of 180 hours in any one quarter of a calendar year. Otherwise, they are classified as co-pilots.

7. Formerly, membership was restricted to white persons, but this provision, which never has been of any practical importance because no colored pilots have ever been employed by an air line, was deleted from the constitution in 1942. The writer has been informed by an official of the National Urban League that a recently organized air express company has made provision to use Negro pilots when it commences operations. Because the Army and Navy trained few Negroes on multiple-engine aircraft, the number of eligibles for commercial aviation is necessarily limited.

extraordinary expenses, provided such assessments are approved by a two-thirds majority vote of the Board of Directors before the expenditures are made.".

No specific requirements for procedure or responsibility in calling strikes or for the payment of strike benefits are found in the ALPA laws. The 1946 strike of pilot crews employed by Transcontinental and Western Air was called by the president of the union after the membership on TWA had given him the authority to order a strike at a date chosen by him. During the strike, the ALPA was reported to have paid benefits of $500 per month to pilots and $200 per month to co-pilots, which were paid for by assessments of as high as $100 per month from pilots and $25 per month from co-pilots employed on other lines.[8] However, TWA pilots interviewed by the writer have denied receiving any benefits. Some assessments were levied; this fact caused some dissension within the union.

Pilots on the various air lines are organized into one or more local councils (local unions), each of which annually elects a local executive council composed of two first pilots and one co-pilot, if the council has jurisdiction over fifteen or less active members in good standing; three first pilots and two co-pilots if its jurisdiction exceeds forty-five. Each local executive council is headed by a chairman and a vice-chairman selected from its first pilot members.

To coördinate their activities, the local councils on each air line are organized into master executive councils. If there is but one local council on an air line, the local executive council automatically comprises the master council; if there are two or three local councils, the master council is composed of all council chairmen, the senior co-pilot councilman in each council, and the vice-chairman from the largest local council; and if there are four or more local councils on an air line, the master council is composed of all the local chairmen, plus the senior co-pilot councilman from the largest of three councils, or fraction thereof, grouped on a geographic basis. Master councils elect their chairman and vice-chairman from among their first pilot members, except where the local and master councils are identical, in which case the local officials are automatically master officials.

As in most labor organizations, the convention of the ALPA is its highest governing body. Conventions are regularly scheduled

8. American Aviation Daily, October 28, 1946, p. 278.

annually, but as they may be postponed by the Board of Directors, they have been held on the average only biennially. Convention delegates are selected as follows: each local council is entitled to one first pilot delegate, "who shall be the Chairman of the Local Executive Council, or an active first pilot member in good standing from such local executive Council . . ." Co-pilot convention delegates are selected on the basis of one for every three local councils, except that each air line is entitled to at least one co-pilot delegate, and that operating divisions of air lines outside the United States are entitled to one co-pilot per local council if the air line divisions "present a marked difference in comparison with conventionally established air lines." The senior co-pilot executive councilman representing the largest council on an air line, or the largest of a group of three councils to be represented by one co-pilot delegate, becomes the convention delegate. If, however, he is unable to attend, he is required to "appoint an active co-pilot member from his Local Council, as his proxy."

Between conventions the Board of Directors has all the authority of a convention, including the right to alter the constitution by majority vote, provided that only three-fourths of the board are present, or vote by proxy or by mail. Since, however, the Board of Directors is composed of the same personnel which comprises the convention delegates[9] — local council chairmen and certain senior co-pilot councilmen — the effect of convening the board is not very different from holding a convention.

Because of the difficulty of convening a Board of Directors scattered all over the globe, ALPA laws formerly permitted the Board to grant authority to the Central Executive Council to act in its stead, which was chosen partially on the basis of the proximity to Chicago, the location of ALPA headquarters. It included the president, secretary and treasurer of the ALPA as ex-officio members, and representatives from local Councils domiciled in and around Chicago, where ALPA headquarters are located. Besides acting on authority from the Board of Directors, the Central Executive Council also served as an advisory board to ALPA officers and as an appellate court for those convicted of violation of ALPA laws by a local executive council. It was abolished by the 1947 convention.

9. With one minor exception — there is no provision for substitutions on the board as in the case of convention delegates.

The present ALPA laws call for twenty national officers —
president, first vice-president, sixteen regional vice-presidents,
secretary and treasurer. Only active first pilots are eligible for
these positions. An executive vice-president was added at the
1947 convention. All officers are nominated and elected at con-
ventions and serve terms of two years, unless recalled by a two-
thirds majority vote acting on petition of 30 per cent of the active
members.

The president is the only full-time officer of the Association,[1]
receiving a salary fixed by each convention in line with earnings
of first pilots at that time.[2] All other Association officers, national
as well as local, retain their piloting positions and are compensated
by the ALPA only for time lost from their jobs on account of
union business. The newly created executive vice-presidency may
be a full-time job, however.

Only first pilots domiciled in or near Chicago are eligible for
the positions of secretary and treasurer, since no others could
perform these duties on a part-time basis. The regional vice-
presidents are elected on a geographic basis, with at least two from
foreign operations.

Although the president is technically subordinate to the Board
of Directors, his constitutional powers, coupled with the vigorous
personality of the incumbent, have placed him in the position of
the dominant, if not sole, policy maker of the organization.
Mr. David L. Behncke was the leading force behind the organiza-
tion of the Association, and he has been its president since its
inception. As the only full-time officer, he has had no rival as to
knowledge of the ALPA's activities and control of its bureaucracy.
Duties performed by full-time elective secretaries or vice-presidents
in other labor organizations are taken care of by his appointees,
whose job tenures and salaries are determined solely by him. He
has negotiated nearly all agreements, and none is effective unless
it bears his signature. Without the approval of "Headquarters"
(which in fact has been Mr. Behncke) no conferences or negotia-
tions on wages or related matters are permitted. Regional vice-
presidents represent the ALPA strictly in accordance with the

1. Prior to 1935, even the presidency was conducted on a part-time basis.
2. It was set originally in 1935 at $7,000 and has since been raised to
conform more nearly with the earnings of the most senior air-line pilots flying
on the more desirable assignments. The 1946 salary of ALPA's president
was $12,000.

president's instructions. The Central Executive Council and the Board of Directors have always been deferential to his wishes.

Moreover, Mr. Behncke has utilized his powers to the full. The ALPA was developed under his leadership, and with little variation its policies and government have reflected his ideas. He monopolizes all communications to the membership. The Association's official publication, "Air Line Pilot," is under his supervision. He has copyrighted the union by-laws and other of its publications. His paper constantly eulogizes him, publicizes his opinions, and gives no space to dissenting views. When occasion arises, an official circular is sent from "Headquarters." No convention proceedings are published, and conventions are closed to the public. Non-official views can be transmitted to the members only through outside sources.

The 1947 convention marked the first formidable challenge to Mr. Behncke's leadership. Charging in speeches to ALPA councils all over the country that the ALPA's "by-laws have been largely designed and written to insure absolute compliance with the president's wishes," and that "Our ALPA . . . has developed into a dictatorship in which the president is the autocratic ruler of the dues-paying member . . .," W. H. Proctor, a senior pilot of American Airlines, ran in opposition for the presidency. He was supported by a large group who felt that the TWA controversy was mishandled, as well as those seeking to democratize the organization, but was defeated 2,870 to 1,241, after an all-night convention session.[3] The election was then made unanimous. The "Air Line Pilot" reported merely that Behncke was unanimously reëlected, and did not mention Proctor's name.

The 1947 convention did approve a plan to establish regional offices. Its sponsors thus hoped to relieve Mr. Behncke and the headquarters of some responsibility and power. The effect is likely to be slight, however, since these regional offices remain under his direction and are apparently to be staffed by persons chosen by him.

Of special importance in maintaining the control pattern within the ALPA is the status of co-pilots. It has already been stated that only first pilots are eligible to serve as national officers or as chairmen or vice-chairmen of local or master executive

3. *Aviation Daily*, February 25, 1947; see below for the TWA case, p. 554; and above, p. 539.

councils. It has also been noted that in all councils and in the convention of the ALPA the co-pilots are represented by a smaller number of delegates than are the pilots. This is true despite the fact that most air lines employ more co-pilots than pilots at any one time. (The number of co-pilots eligible to ALPA membership is, however, made smaller by the fact that one year's experience is a prerequisite to membership.) The co-pilots are further subordinated by provisions which grant first pilots in all governmental bodies of the Association one vote for each member they represent, while co-pilots have but one vote for each two members which they represent.[4]

All forms of union security and check-off are forbidden under the Railway Labor Act. Nevertheless, both ALPA and company officials agree that over 90 per cent of all scheduled air line pilots are members of the ALPA. This can be explained to a considerable extent by the tremendous accomplishments of the union for the benefit of air-line pilots, but that does not tell the whole story. The question might, for example, be raised as to why co-pilots are so quick to join an organization which is dominated by the pilots.

The answer is not difficult to discover. All co-pilots are desirous of becoming pilots, because of the greatly increased remuneration which results from such promotion. The air lines in the interests of safety and efficiency regard the first pilot and co-pilot as a team. No first pilot is forced to fly with any co-pilot to whom the first pilot objects. ALPA first pilot members make it a practice of pointing to the advantages of joining the organization, and it is generally understood that if a co-pilot does not respond, fault is soon found with his efforts by the first pilot to whom he is assigned. As the co-pilot must fly with various first pilots as part of his training, and as he must have the approval of each first pilot in order to fly with him, he soon understands that membership in the organization is advantageous.

Affecting not only membership of co-pilots but of pilots also, is the fact that check pilots are generally drawn from senior active pilots (and hence ALPA members). The check pilot regularly examines the work of pilots and co-pilots, in order to insure that

4. Originally co-pilots were not eligible to membership. Later they were admitted without voice or vote. Present arrangements were worked out in the 1936 and 1938 conventions.

their abilities conform to the minimum standards required by the Civil Air Regulations. Check pilots are permitted by ALPA laws to retain active membership, except that they are not eligible for union office. Their sympathy for the ALPA is such that one frequently hears tales of how they actively "encourage" union membership.[5] In addition chief pilots, who are in charge of pilots, are former ALPA members on most lines. They, too, are often sympathetic to their old ties.

Other reasons why co-pilots and pilots usually join the ALPA soon after they become eligible for membership include the fact that dues and initiation fees are considerably lower for pilots than for co-pilots; that co-pilots have been promoted on average thus far after only two and one-half years' experience, and so long as this average is maintained, the status of co-pilots in the ALPA is not of major importance to that group; and the provision (Article III, Section 3 (b)) of the ALPA constitution that anyone who fails to join the Association when he becomes eligible "shall be deemed to have accepted the benefits of the Association without assuming any of the obligations thereof and as a prerequisite of his being accepted as an active member, he shall be required to pay, in addition to the regular initiation fee, all of the annual dues which would have accrued to him as an active member of the Association beginning one year after such applicant first became eligible for membership."

In order to provide for the affiliation of non-pilot or foreign groups, Article I, Section 10 of the ALPA Laws provides that the president, subject to ratification or change by the Board of Directors, may grant charters of affiliation; but since the charters of affiliation require that the affiliated parties adhere to ALPA laws without also providing for any rights to the affiliated parties to assist in the making of these laws, and since charters of affiliation may be revoked at the pleasure of the ALPA, potential affiliates have been more alienated by these rules than induced to affiliate. In fact, the only affiliated groups are now the British and Canadian

5. The following story was related to the writer in November, 1946. Several co-pilots on a large air line held membership in the Military Pilots Association. Check pilots on that line refused to certify these co-pilots for first pilot positions until they joined the ALPA and presented certificates from the secretary of the MPA affirming the fact that they were no longer members of that organization.

pilot associations which, by special arrangements, have mutual affiliation agreements without such onerous conditions, and the stewardesses who have been organized by ALPA.

METHODS OF SALARY PAYMENT AND EARNINGS, 1925-1945

The Pre-NRA Period. The complicated method of pilot remuneration now in effect derives from the system developed by the Post Office in the early 'twenties, when flying was extremely hazardous and uncertain. In 1925 the Post Office paid mail pilots a minimum base salary of $2,000 for daylight flying, which progressed to $3,600 at the rate of $100 per annum for each 500 hours in the air. A night bonus of $400 per annum was paid for day-and-night flying, where the latter did not exceed one-third of mileage; of $600 where night flying exceeded one-third of mileage; and of $800 where schedules called for night flying.[6]

In addition to base salary, the Post Office paid mileage allowances of varying rates, depending upon whether day or night flying was utilized, and upon the relative difficulty of the terrain. Thus, for day flying, pilots were paid five, six, and seven cents per mile flown. The five-cent rate was in effect between Cleveland, Ohio, and Cheyenne, Wyoming; the six-cent rate, between Cheyenne and Reno, Nevada; and the seven-cent rate between New York and Cleveland and between Reno and San Francisco. For night flying, the mileage pay was doubled and pilots received 10, 12 and 14 cents, respectively. Monthly earnings of mail pilots in 1925, the last year of government operation, averaged $563.33.

Private contractors who took over government routes in 1926 generally continued for several years the salary schedules inaugurated by the Post Office Department. Those who developed new routes, however, tended to vary methods of payment, and the depression commencing in 1931, plus the introduction of faster planes, resulted in a general re-examination of pilot pay scales. Nevertheless, a survey by the United States Bureau of Labor Statistics in the fall of 1931[7] found that monthly base pay plus mileage remained the prevailing method of pilot remuneration. Most companies still paid higher mileage rates for night flying,

6. The discussion of the Pre-NRA period is based upon Federal Coördinator of Transportation, op. cit., pp. 34–51.

7. "Wages and Hours of Labor in Air Transportation, 1931," Bulletin No. 575, U. S. Bureau of Labor Statistics, 1933.

but night mileage rates were no longer twice as high as day rates. Mileage pay likewise continued to vary with terrain. On the other hand. in 1931 carriers employing 25 per cent of the pilots paid a flat monthly salary without regard to mileage flown, and flat monthly salaries were the usual form of compensation for co-pilots.

During the three months prior to the 1931 survey of the Bureau of Labor Statistics, general pay reductions throughout the industry took place, amounting to as high as one-third to one-half of base pay, 10 to 36 per cent in mileage rates. Improved, faster planes kept earnings up, however, pilots in October, 1931, averaging $569.49, as compared with $563.33 six years previously. Co-pilots averaged $227.89 in October, 1931.

By July, 1933, payment by base plus hours flown had replaced base plus mileage as the dominant method in the industry. In addition, some carriers continued paying flat salaries. Hourly pay varied, according to company, terrain, and type of airplane, from $3 to $5.50 per flight-hour for day flying, and from $4.80 to $8.25 for night flying. A field survey of 462 pilots in July, 1933, found them earning an average of $621.33 monthly. Most of the increase over the 1931 average was accounted for by longer hours, but hourly earnings were 2.4 per cent greater in 1933 than.two years before.

Effective October 1, 1933, the one remaining large company which had not shifted from mileage to hourly remuneration announced its intention of so doing. Coincidentally, the "Big Five" carriers[8] announced that as of that date, when higher-speed ships were introduced, a new pilot pay scale would be adopted, using base plus hourly pay and, except for one carrier, abandoning differentials for difficult terrain. A dispute with the ALPA led to the assumption of jurisdiction by the NRA National Labor Board, and Decision 83 resulted.

Decision 83. This provided for annual base pay of $1600–$3000, increasing from minimum to maximum annually by $200 increments on the basis of seniority; for hourly pay varying with speed, with a 50 per cent differential for night flying (i.e., $4, $4.20, $4.60, $4.80 and $5 for day flying and $6, $6.30, $6.60, $6.90, $7.20 and $7.50 for night flying at hourly speeds of under 125 miles per hour, 125 miles, 140 miles, 155 miles, 175 miles, and 200 miles or

8. American, United, Transcontinental & Western, Eastern, and Western. These are not the five largest today.

more, respectively); and for mileage pay at monthly mileages of under 10,000, 10,000 to 11,999 miles and 12,000$\frac{1}{4}$miles and more, respectively, of two cents; one and one-half cents, and one cent per mile for all miles flown at an hourly speed in excess of 100 miles per hour.

DIAGRAM I
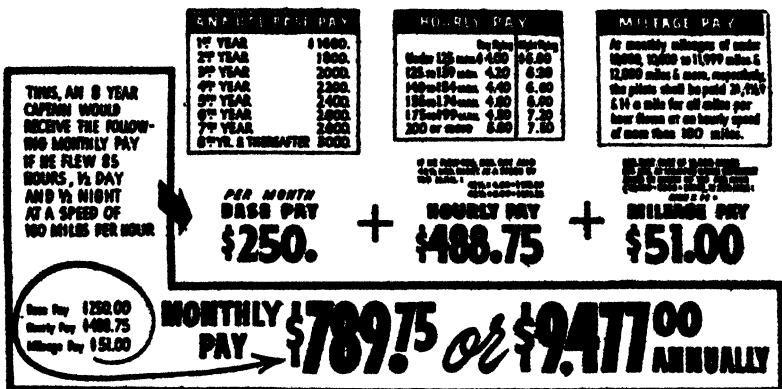
SOURCE: TWA EMERGENCY BOARD, CARRIERS' EXHIBIT No. 16

Diagram I explains how, on the basis of Decision 83, the monthly salary of an eight-year captain (pilot), who flies the maximum allowable of 85 hours, one-half during the day and one-half during the night, at an average speed of 160 miles per hour (the DC–3 average) is computed. To the base salary of $250 is added $195.50 hourly pay for day flying (42.5 hours at $4.60), $293.25 hourly pay for night flying (42.5 hours at $6.90), and $51 mileage pay (85 hours at 160 m.p.h. equals 13,600; deduct 85 x 100 to get mileage flown at speed greater than 100 m.p.h.; multiply result, 5,100, by $.01, since monthly mileage exceeded 12,000: result equals $51).

The bases for determining flight hour pay vary in different ALPA carrier-agreements. Most contracts provide for pay on the basis of the scheduled "block to block"[9] time for scheduled flights,

9. Block to block is the term used to mean the time from which the blocks are removed from the plane's wheels, preparatory to take-off, until the pilot cuts the plane's engines after landing at his destination.

and actual block to block time for non-scheduled flights. Others, however, require actual block to block time on all flights. On scheduled flights governed by the former provision, a pilot is paid his scheduled time if he completes his flight faster than schedule; but if he is held up by air traffic delays more than a specified minimum (usually fifteen minutes), he is paid for all time over that minimum.

ALPA-carrier contracts are even less uniform in stipulating the speed to be used in determining the mileage component for each type of plane. For example, for the twin-engine DC-3, some provide for scheduled speed, but not less than 160 miles per hour; others for actual monthly or annual averages. The agreement with Delta Air Lines has no special provisions; others use a speed minimum of 155 miles per hour or a bracket of 155-175. Despite the great variation, the results are usually not very different.

Decision 83, as written into law, provides not only the highest minimum wage law in the land, but also one which increases with advancing technology. Table I, which compares average monthly earnings for pilots for selected years since 1925, does not show the effects on these earnings of Decision 83 to a great extent, because the continued expansion of the industry, and its pilot force, reduced the percentage of senior pilots in the more recent years.

The practical effect of Decision 83 on pilot earnings is better

TABLE I

AVERAGE MONTHLY EARNINGS OF PILOTS EMPLOYED BY SCHEDULED AIRLINES
IN DOMESTIC SERVICE, SELECTED YEARS, 1925-1945

Year	Average Pilot Monthly Earnings
1925	$563.33
1931	569.49
1933	621.33
1935[1]	668.49
1937	653.72
1939	704.19
1941	702.44
1945[2]	718.78

Source: Federal Coördinator of Transportation, op. cit., pp. 34-37, for 1925-35; Records of the Airline Negotiating Conference for 1937-45.

1. Excludes non-mail companies, who were not covered by Decision 83. Their inclusion would drop 1935 average to $663.93.
2. Data for 1945 have downward bias because all Stratoliner and four-engine pilots, and all pilots on army or navy contract flying are excluded. These positions were held by senior pilots to a large extent.

demonstrated in other ways. Diagram II illustrates how pilot earnings would automatically increase under Decision 83 rates as the air lines substitute faster four-engine equipment — DC–4 (Sky-

DIAGRAM II
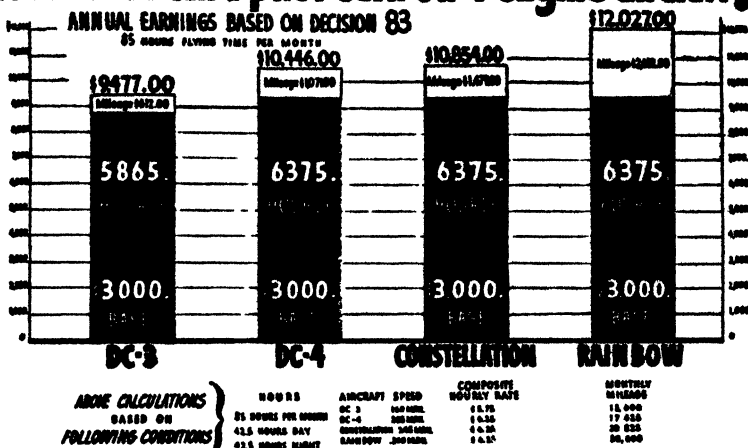
SOURCE: TWA Emergency Board, Carriers' Exhibit No. 18

master), Constellation, and Rainbow — for the twin-engine DC–3, the prewar standard carrier. Fifty-seven pilots employed by American Airlines under Decision 83 rates received average annual earnings of $6,257.35 in 1935, $9,097.39 in 1941, and $10,364.45 in 1945.[1]

Terrain Pay. Besides providing for the base, hourly, and mileage pay formula, Decision 83 also required that "terrain pay" — i.e., extra compensation for flying over hazardous terrain — be maintained. As has already been noted, such extra compensation was originally paid by the Post Office Department between New York and Cleveland, Cheyenne and Reno, and Reno and San Francisco. The reasons for this have been well stated by Mr. John A. Herlihy, vice-president of United Airlines, who flew the New York-Cleveland run as a pilot in the early thirties:[2]

1. Transcontinental & Western Air et al; Before Presidential Emergency Board, 1946, Carriers' Exhibit No. 24. The 1945 data include some extra compensation, but the total upward bias is slight.
2. Ibid; testimony, vol. 3, p. 1678. (Hereafter cited as "TWA Emergency Board.")

New York to Cleveland on the Trans-Continental was set up as a terrain
sector way back, and has been carried forward ever since. Originally, when
terrain pay was first put in, and probably with some justification, there were
a number of accidents in the New York to Cleveland area. The weather
conditions were such that it was difficult to fly the route on contact flying by
just looking out of the cockpit. Furthermore, the planes they operated in
those days were of such short fuel range that it was necessary to get into
Bellefonte, Pennsylvania about half way between New York and Cleveland
to get down into a narrow valley between a couple of mountain ridges, in
order to get gasoline to proceed west with the mail.

Another thing, the planes were subject to icing conditions and rather
dangerously so. Furthermore, they lacked the airway aids, radio ranges
and two-way radio communications, to keep the pilots advised of weather
conditions ahead, and so forth, and as a result there were quite a number of
accidents on the New York-Cleveland run, and that was the basis for the
higher pay on that run. It was strictly a hazard pay.

At the time when Decision 83 was promulgated, all air lines
but two, United and National Parks Airlines, which Western has
since absorbed, had abandoned terrain pay, and at least one of
these had announced its intention of so doing. Decision 83 has
compelled these air lines to continue terrain pay bonuses of one
dollar per flight hour for day flying and two dollars for night flying
over certain areas. In addition, Northwest Airlines, which is under
no legal obligation to pay terrain bonuses, has been induced to do
so by the ALPA in collective bargaining negotiations.

Some of the original terrain runs actually no longer exist. For
example, United Air Lines has consolidated its operating divisions,
in view of modern plane mileage capacity, so that pilots who
formerly flew from Chicago to Cleveland, or from Cleveland to
Washington, New York or Boston, now fly from Chicago to one of
the Atlantic seaboard cities. To facilitate the computation of
pilot pay, the ALPA and United Air Lines have agreed on the
"composite pay rates," which are the average of the terrain and
non-terrain hourly rates on pilot runs, weighted in proportion to
the mileage of the terrain sector to the mileage of the non-terrain
sector, plus certain additions to terrain areas. Northwest Airlines
was persuaded to adopt a composite rate for its recently awarded
Detroit-New York run.

The conditions which prompted the adoption of terrain pay
can hardly be said to exist today, except perhaps in some very
unusual circumstances. Improvements in planes, radio navigation
aids, ground facilities and a host of other technological and scien-

tific advances have largely outmoded 1933 concepts of hazardous terrain. Moreover, the perpetuation of terrain pay has created differentials which are without real basis. For example, on the New York-Chicago run, three air lines, American, Pennsylvania-Central, and TWA, pay no terrain bonus, but United does. Western has to pay terrain bonuses on some routes, but is not compelled to do so on its Denver to Los Angeles route, where planes fly at 15,000 feet, the highest flying altitude over the Rocky Mountains. Similar inequities may be found on all routes on which terrain pay exists. It is difficult to disagree with the conclusion of the TWA emergency board: "The requirement that some companies shall pay for hazardous terrain because they were paying it in 1934, while others are not so required, works unfairly for both carriers and pilots."[3]

Co-pilots' Pay. The only provision made for co-pilots in Decision 83 was that "the differentials existing on October 1, 1933 . . . shall be maintained." At that time co-pilots were paid flat monthly salaries almost without exception, and were generally employed at an initial rate of $190 per month, which was increased to $210 per month after six months and $225 per month after one year. Co-pilots who qualified as first pilots but who did not have regular runs (i.e., were "reserve pilots") were then, as now, paid first-pilot rates for service performed in that capacity. In July, 1933, co-pilots' average earnings were $231.13, as compared with $227.89 in 1931.[4]

Since the issuance of Decision 83, co-pilot salaries have been increased on all the scheduled airlines, either by unilateral carrier action or by collective bargaining. The agreement effective October 21, 1944, between Transcontinental & Western, from which there is little variation in the industry, except for four-engine equipment, provides for a starting co-pilot rate in domestic service of $220 per month, with automatic increases of $20 per month every six months till a maximum of $380 is reached after four years. Since, under the present conditions of rapid growth in the industry, co-pilots may expect promotion to first pilot after two and one-half years of service, few ever achieve maximum co-pilot pay.

International Service Pay. The 1934 and 1935 air-mail laws,

3. TWA Emergency Board, Report to the President, pp. 31–32.
4. Federal Co-ördinator of Transportation, op. cit., pp. 47–51.

which incorporated Decision 83, made no mention of a distinction between domestic and international service. The Civil Aeronautics Act of 1938, however, did so, providing that international air carriers "shall maintain rates of compensation" for pilots and co-pilots "the minimum of which shall not be less, upon an annual basis, than the compensation required to be paid under said Decision 83 for comparable service" in domestic operations.

The pioneer international air carriers, Pan American Airways, Inc., has always paid its pilots a flat monthly salary. The ALPA-Pan American agreement, which was dated June 16, 1945 — the first between the parties — does, however, provide for bonuses for night flying on an hourly basis and for overtime, in addition to flat monthly salaries graduated on the basis of seniority.

The monthly salary scales of Pan American pilots and co-pilots provide for two categories of pilots and co-pilots; master and regular. The former scale was first evolved for flying boats and was later used for four-engine equipment; the latter is the two-engine scale. Monthly salaries for pilots have varied from $627 to $830 for the regular group, and from $953 to $1,060 for the master group; for co-pilots the ranges are $240–$470, and $340–$570. Pilots reach maximum rates after eight years of service; co-pilots after four years.

In addition to these monthly salaries, Pan American's "regular" pilots receive $2.50 for each hour flown at night; and each pilot whose flight duty exceeds 340 hours in any quarter-year or 900 hours in any calendar year, receives overtime determined by multiplying his total overtime hours by 1/81.8 times his average monthly salary.[5]

Pan American-Grace Airways (Panagra), which operates in South America north to the Canal Zone, pays, pursuant to its agreement with ALPA, flat salaries ranging for co-pilots from $275 to $450 per month, for pilots, from $700 to $875 per month, after eight years of service.[6] Panagra's rates for the DC-3 thus slightly exceed pilot earnings in domestic service operating under the terms of Decision 83 (see Diagram I, above).

5. Agreement between Pan American Airways, Inc. and . . . Air Line Pilots Association, June 16, 1945, pp. 11–13.
6. Supplemental Agreement between Pan American-Grace Airways, Inc. and . . . Air Line Pilots Association International, November 15, 1943, pp. 2–3. This schedule contained increases over the first (1941) agreement.

The second transoceanic carrier to commence operations, American Export Airlines (now American Overseas Airlines), paid by agreement with the ALPA, effective December 1, 1944, a monthly salary schedule similar to that of Pan American. Instead of night or overtime bonuses, however, American Export's agreement provided for a monthly bonus of $50 or 15 per cent of salary, whichever was higher, to all pilots and co-pilots stationed outside the continental limits of the United States and other than in the country of his nationality; in addition, in March, 1946, ALPA executed a supplemental agreement with the carrier, providing for hourly flying pay, day or night, of two dollars per hour for co-pilots.[7] This, however, applied only to a military contract operation, and hence is not now in effect.

During the war a number of domestic carriers undertook overseas operations for the Army Transport Command and the Naval Air Transport Service. Some carriers, e.g. Transcontinental and Western, put a rate into effect on army orders; others, e.g. Northeast, with army permission, worked out an agreement with the ALPA, whose members supplied the nucleus of the armed forces air transport personnel. These monthly rates varied from $955 to $1,150 for pilots and from $450 to $637.13 for co-pilots, both on an eighty-five hour basis.[8] The high war-time salaries reflect the many clements of war risk which were necessarily involved. It is noteworthy that Northeast, which had the highest rate, flew two-engine planes through the treacherous weather and territory of Northern Canada and Greenland.

The Stratoliner Arbitrations. In 1940, Transcontinental and Western Air introduced the four-engine Boeing Stratoliner, the first important departure from the standard DC–3 in several years. Although Decision 83 would automatically have provided for increased compensation for flying the faster Stratoliner, the ALPA requested extra compensation on the ground that it was a new, heavier, faster, and more productive craft. The issue went to arbitration, and the board of arbitration, without stating reasons, awarded a bonus above Decision 83 rates for flying Stratoliners of 80 cents per hour for day flying and $1.20 per hour

---

7. Agreement between American Export Airlines, Inc. and ... ALPA, December 1, 1944, pp. 3–4; Supplemental Agreement between American Overseas Airlines, Inc. and ALPA, March 21, 1946, pp. 2–3.

8. TWA Emergency Board, Report, p. 19; ALPA-Pan American-Grace contract, November 15, 1943; pp. 8–9. Actual earnings exceeded the rates because extra hours were frequently flown.

for night flying for pilots and a $20 per month bonus for co-pilots. In 1945 a new model of the Stratoliner was introduced, and again the same controversy arose. This also went to arbitration, and once more without stating reasons an arbitration board awarded pilots a further increase on the new Stratoliner of 80 cents per hour for day flying and $1.20 for night flying, and awarded co-pilots an additional monthly payment of $20.[9]

The Stratoliner cases marked the adoption by the ALPA of a new wage philosophy. Abandoning its doctrine of a sliding scale (i.e. Decision 83), it now sought to win new and higher rates for each new type of craft which was introduced. The climax of this new policy came with the introduction of larger four-engine planes in late 1945.

The 1945–1947 Four–Engine Dispute

During the war a number of larger, faster, four-engine airplanes were developed, but their use was confined to military operations. With the end of the war several air lines made plans to introduce this new equipment, especially the Douglas DC–4 (or Skymaster) and the Lockheed Constellation. Other new planes then in progress included the Boeing Strato-Cruiser (a commercial development of the B–29 bomber), the Douglas DC–6, and the Republic Rainbow. As in the case of the earlier Stratoliner, Transcontinental and Western Air pioneered in the development of the four-engine equipment. Since TWA was very anxious to put the DC–4's and Constellations into operation at an early date, especially in its recently authorized international operations, it approached the ALPA to discuss rates for four-engine planes and for international operations in July, 1945.

After some preliminary conferences, negotiations commenced in September, 1945. Offers of TWA, first to pay Pan American rates, and later bonuses above Decision 83 rates averaging three dollars per hour for DC–4's and four dollars per hour for Constellations,[1] were rejected by the ALPA, which literally refused to discuss

9. For copies of these awards, see TWA Emergency Board, ALPA Exhibit No. 49, and Carriers Exhibit No. 39. The effect of the Stratoliner arbitrations on pilot salaries is as follows: a pilot with eight years seniority flying 85 hours per month, one-half day and one-half night, at the average Stratoliner speed of 190 m.p.h., would receive, under Decision 83, $836.50. Under the first Stratoliner award, he would receive an additional $85; under the second, an additional $170.

1. This section is based largely on the record in the TWA emergency board case. The effect of the TWA bonus offer would have been to pay $4,080.

any rates except for domestic DC–4 operation, and which meanwhile instructed its members to refuse to bid on DC–4's or Constellations for either domestic or international service. As a result, TWA's new fleet of planes were grounded, inauguration of its international service was postponed from September, 1945, to February, 1946, and Pan American placed Constellations in international service before it was able to do so.

While TWA negotiations were in progress, the ALPA was making demands on other air lines which were contemplating using four-engine equipment. On these air lines the ALPA likewise instructed its members to refuse to bid on four-engine equipment, on the ground that there was no rate covering their operation, although, as the carriers pointed out, this reasoning ignored the existence of Decision 83 and relevant provisions of ALPA-carrier agreements.

These protracted negotiations and the fear on the part of the air lines that they would be repeated each time new equipment was introduced, led in December, 1945, to their forming the Airlines Negotiating Committee (now Conference) for the purpose of handling jointly negotiations on four-engine equipment.[2] Mr. Behncke refused to recognize the Committee, terming it "illegal" and announcing his determination to deal with the carriers only on an individual basis. Even after the carriers' committee had received full sanction of the Civil Aeronautics Board,[3] the ALPA refused to recognize it "as such," either for the purpose of multi-

above the annual Decision 83 rate for Constellations, and $5,610 for DC–4's, on the basis of 85 hours per month, flown one-half day and one-half night. When it is noted that an eight-year pilot, under Decision 83, at average speeds of 245 and 205 m.p.h., respectively, would receive $10,854 per month on the Constellation and $10,446 on the DC–4, one can perceive that this was a generous offer, to say the least.

2. Originally the Airlines Negotiating Committee was composed of thirteen carriers which planned to use four-engine equipment, and its purpose was restricted to dealing with ALPA on that issue. Later, it was reconstituted as a permanent conference to handle collective bargaining with ALPA for its members, and with any other group of employees if any carrier member so desired. The membership of the conference includes all major domestic and international certificated carriers except Colonial, Hawaiian, Alaska, Caribbean-Atlantic, Pan American, Pan American-Grace, and some feeder lines.

3. All agreements among certificated air carriers are made subject to CAB approval by the Act of 1938. The ALPA vigorously opposed such approval in two hearings before the CAB.

carrier bargaining or as the bargaining representative of the individual air lines, despite the fact that multi-carrier bargaining was by no means novel in the industry,[4] that President Behncke had in previous years been the leading advocate of uniform wages and working conditions among the air lines[5], and that the Railway Labor Act purports to guarantee to employers, as well as to employees, the right to determine their bargaining agents.[6]

Despite ALPA's refusal to recognize it "as such," the Airlines Negotiating Committee took over negotiations in all four-engine cases for carriers which it represented. After TWA officials, who are also licensed pilots, had flown the testing runs, TWA, in February, 1946, pursuant to its agreement with ALPA, assigned junior pilots to its international division.[7] They accepted "under protest" and henceforth the four-engine equipment on TWA and other air lines was bid on by pilots, usually after a statement that it was being done under protest.

Negotiations dragged on for five more months, with the ALPA never explicitly stating its demands, and with fruitless attempts by the National Mediation Board to effect a settlement by mediation or voluntary arbitration. In April, 1946, the pilots on TWA voted to strike, and in May, 1946, to avert a stoppage, President Truman appointed an emergency board to hear not only the TWA case, but those involving twelve other air lines as well.

The board, composed of Judge George E. Bushnell, Dr. William M. Leiserson, and Dr. John A. Lapp, sat from May 17 to June 19, and heard testimony and argument consisting of 3,495 pages and 266 exhibits. For the first time, the ALPA made explicit wage demands, which ranged up to approximately $20,000 per year for operating Constellations in international service. The

4. The ALPA instituted such bargaining itself in 1933 in the case which resulted in Decision 83. In 1942, ALPA and all domestic certificated carriers signed an industry-wide agreement extending maximum hours from 85 to 100 dollars lance with the temporary legal extension during the war. (See

tions,[1] were re testimony before various Congressional Committees, as

9. For copVA Emergency Board Hearings, Vol. I, pp. 198-202.
Exhibit No. 49,2, Third, of the Railway Labor Act, as amended, states: arbitrations on p for the purposes of this Act shall be designated by the flying 85 hours p without interference, influence or coercion by either party Stratoliner speed m of representatives by the other . . ." Section 2 (1) sets Under the first St of the Act, among others, as "to provide for the complete the second, an addriers and of employers in the matter of self-organisation."

1. This sectitp. 565, for a discussion of seniority.
board case. The of

Airlines Negotiating Committee, which had originally gone on record in favor of "reasonable differentials in compensation . . . for the operation of larger equipment" maintained that, after studying carefully Decision 83, it could only conclude that such differentials were provided by Decision 83. In addition, both parties set forth their views on various ALPA proposed fringe issues and rule changes, and on the question of whether the board should concern itself with only the TWA case, as the ALPA argued, or with all the air line cases treated as one, as the carriers maintained.

The emergency board's report, issued on July 8, 1946, rejected the ALPA proposal of "a new bargain as to base pay, hourly pay, and mileage pay for each new type of plane that is put into service," but it likewise did not accept the carriers' contention that Decision 83 was adequate. Instead, it proposed supplementing the hourly and mileage pay components of Decision 83 as follows: (1) extend the hourly pay brackets from the 200-mile and over maximum limit by 20-cent increments for day flying and 30-cent increments for night flying for each additional 25 miles flown per hour;[8] (2) substitute the rate of 1.5 cents per mile for that of 1 cent per mile in figuring the mileage component at monthly mileages of 12,000 or more flown at speeds in excess of one hundred miles per hour; and (3) extend Decision 83, as thus supplemented, to international service, but increase pilots' base pay in international service by $750 per year in all seniority pay brackets.[9]

The emergency board was convinced that the differential between pilot and co-pilot pay was "too great, particularly in international operations." It therefore recommended that in domestic service co-pilots should receive in addition to their base pay "mileage pay at the rate of one cent for monthly mileages of 12,000 miles or more flown at speeds in excess of one hundred miles

8. The following table illustrates how this recommendation would add to the hourly pay component of Decision 83:

M.P.H.	Rates	
	Day	Night
200–225............	$5.00	$7.50
225–250............	5.20	7.80
250–275...........	5.40	8.10
275–300...........	5.60	8.40
Etc.		

9. TWA Emergency Board, Report, pp. 30–32.

an hour." The board further recommended that on international operations "co-pilots should receive, in addition to base pay, hourly flying pay at a composite rate of $1 an hour for day- and night-time flying. They should also receive mileage pay at the rate of one cent for monthly mileages of 12,000 miles or more flown at speeds in excess of one hundred miles an hour. Co-pilots in international service should receive $25 a month additional base pay if they qualify as navigators."[1] The emergency board also made recommendations for rules in international service.[2]

The basic wage and rule recommendations of the emergency board were applied only to the TWA case. Insofar as the other twelve were concerned, the board recommended that they be referred back to the parties for settlement by agreement on the basis of TWA recommendations. The board declined to accept the ALPA position that these twelve cases were not properly before it, pointing out that its jurisdiction was governed by the executive order which created it. On the other hand, the board refused to order industry-wide collective bargaining, because it felt that industry-wide bargaining is most successful where the parties agree to it, not where it is imposed unilaterally or by a government agency. It did, however, declare that "the right of each of the air line companies to designate the (Airlines Negotiating) Committee as its authorized bargaining representative cannot justifiably be questioned by the (pilots') Association.[3]

After the issuance of the board's recommendation, the Airlines Negotiating Committee, as well as various officials of the airlines, made several vain attempts to arrange a meeting with the ALPA in order to work out agreements based on the recommendation, but President Behncke continued to decline recognition of

1. Ibid, p. 33.
2. See below, pp. 564–568, for a discussion of rules.
3. Report, pp. 37–38. The effect of the emergency board report on monthly rates is shown below for an eight-year pilot, flying Constellations at 245 m.p.h., one-half day and one-half night, in domestic service (international service received an additional $62.50 per month):

	Decision 83 Rates	Emergency Board Rates
Base Pay	$250.00	$250.00
Hourly Pay	531.25	55250
Mileage Pay	123.25	184.88
Total Monthly Pay	$904.50 ·	$987.38

the Airlines Negotiating Committee as the carriers' bargaining agent. Moreover, he took the position that no agreement was possible until the board had interpreted several of its recommendations. Finally, one month after the board's recommendation had been issued and hence when the period ended during which unilateral action for matters in dispute is prohibited by the Railway Labor Act, TWA put the board's recommendations into effect. The result was to grant substantial increases to most of TWA's piloting crews, but to reduce the rates of a few who formerly flew under TWA's Army Transport Command contract.[4] Moreover, for international flying the emergency board rates were considerably below ALPA's demands and previous TWA offers, including the offer to pay existing Pan American rates.[5]

Within a month after TWA acted thus unilaterally, all but one of the other twelve airlines involved in the dispute also put the new rates into effect.[6] The Airlines Negotiating Committee, however, continued its attempts to arrange a meeting with the ALPA and work out agreements, but without success. Then, after thus declining to negotiate for almost three months, Mr. Behncke called President Jack Frye of TWA late Friday afternoon, October 18, and asked to see him the following morning, indicating that a strike would result if the conference could not be arranged. Mr. Frye agreed, and at this conference the ALPA leader demanded maximum rates in excess of $1,250 per month for Constellations and $1,000 per month for DC-4's, both for domestic operations, with an overseas differential to be determined later and an answer at the close of business on the following day, which was Sunday.

4. When TWA assigned pilots to its international division, it paid them rates based on Decision 83, or on their previous six months average earnings, whichever was higher, so that it could not be accused of altering rates unfavorably during negotiations, and thus acting contrary to the Railway Labor Act's procedure. Included in the group paid average earnings were a few who had been receiving $1,100 per month under Army Transport Command work. They suffered some reduction when the emergency board recommendations were put into effect, but they included less than ten per cent of the total affected personnel.

5. Under conditions as specified in Note 9, p. 554, TWA international Constellation senior pilots would be paid $1,049.88 per month under the emergency board rates. For this, ALPA had demanded approximately $1,700, TWA had offered to pay $1,244.50 and Pan American was paying $1,060.

6. Northwest was the exception, because ALPA had indicated willingness to negotiate with it soon after TWA acted unilaterally.

TWA officials replied within the ultimatum time, but instead of agreeing to the rates, suggested a conference, noting that they had been trying to obtain one for some time. Whereupon TWA pilots struck at 4:30 a.m. Monday, October 21, 1946.

Conferences were soon arranged by the National Mediation Board, and a tentative agreement to arbitrate was worked out. On November 4, the ALPA announced that it had "officially accepted" an arbitration offer, and stated further that it was forced to strike only because of the "monopolistic air lines negotiating trust," which was trying to destroy organized labor on the air lines. Nevertheless, the arbitration agreement was signed, on November 6 by TWA, but not until November 15 by ALPA and then only after several changes had been made in the wording of the agreement.[7]

In accordance with the agreement, a three-man tripartite arbitration board re-heard the issues already passed upon by the emergency board to determine, realistically speaking, how much more than the emergency board recommended should be awarded to the TWA pilot crews. The arbitrators unanimously granted pilots flying the DC-4's and Constellations in domestic service base pay increases of $50 per month and slight revisions in the computation of the mileage formula recommended by the emergency board, which would yield pilots $7 to $10 per month. The emergency board's recommendations for hourly pay were left intact.

For international service, the arbitrators abandoned the Decision 83 formula in favor of flat rates plus overtime, similar to the Pan American scale. Those pilots flying DC-4's and Constellations in international service received $925 to $1,100 per month, with a graded scale advancing with length of service from one to eight years, and in addition received $13.50 for each hour flown in excess of nine hundred hours in any calendar year. Since the Civil Air Regulations limit international flying to one thousand hours per year, and since few pilots fly much in excess of nine hundred hours, the overtime will not result in any substantial payments.

For domestic service, co-pilots flying DC-4's were awarded rates from $260 to $420 per month, and those flying Constellations, $280 to $460 per month, both increasing from minimum to maxi-

7. Based on official documents and correspondence in the case.

mum in a four-year period. In international service, co-pilots flying both DC-4's and Constellations were awarded $290 to $520 per month, with the same period for upgrading, plus $50 per month to all co-pilots qualified as celestial navigators. As in the case of the emergency board, the arbitrators provided for various rule changes which are discussed below.[8]

Thus, after more than a year the four-engine dispute was brought to a close on TWA. Despite the long and costly struggle, however, TWA pilots in international service ended up with rates virtually the same as those paid by Pan American, which the company offered to pay in the first place, and with considerably less, both in domestic and international, than the later TWA offer of three-dollar and four-dollar per hour bonuses for CD-4's and Constellations, respectively.[9]

The outcome of the TWA dispute did not serve as a model for future negotiations on other airlines. New contracts negotiated since then have all called for higher rates, e.g., as much as $148.75 per month for Eastern Air Lines' Constellations, and approximately as much for the new DC-6 on United and American. Moreover, most new contracts contain a new factor, plane weight, which is figured into the rates. And as if to stress the general dissatisfaction with the TWA outcome, that company and ALPA have signed a new agreement setting aside the arbitration award for co-pilots and have returned to single co-pilot rates regardless of equipment flown in domestic service.

Another interesting aspect of the four-engine dispute was the minor revolt against ALPA leadership. President Behncke, who assumed full charge of the emergency board hearings and of the strike (as well as of all previous ALPA activities), was relegated to the background in the arbitration hearings, which were handled by an attorney selected by the TWA pilots. As noted, however, Behncke survived the 1947 convention without serious loss of power.

8. Based on the arbitration report rendered January 22, 1947. All increases were made retroactive to November 15, 1946, the date of the arbitration agreement.

9. See data in Note 5, p. 559. An eight-year pilot in domestic service under conditions there described will net approximately $60 per month above the emergency board rates. International rates of $1,100 per month plus overtime are equivalent to Pan American's $1,060.

Finally, the four-engine dispute serves once more as a reminder that the emergency board "fact finding" procedure in the Railway Labor Act is not only far from a "model" for more far-reaching legislation, but that it has not settled one major dispute since 1941 in either the railway or the air lines industries.[1]

HOURS OF WORK AND MILEAGE LIMITATIONS

In 1931, the Department of Commerce, using for the first time its authority under the Air Commerce Act of 1926, set maximum hours for pilots of 110 hours per month, not more than 30 hours of which were to be scheduled in any seven days nor eight hours in any 27-hour period.   Twenty hours of relief from flight duty were required within each seven-day period.[2]   Particularly during the worst depression years, these regulations were not well heeded or enforced.   This led to demands by the ALPA to reduce pilot hours to 80 per month, which Decision 83 rejected in favor of the figure of 85, as recommended in 1933 by the Aero Medical Association.

No maximum was established for co-pilots' flight-hours until October 1, 1934, when the Department of Commerce prescribed a limit of 100 per month, without, however, any supplementary yearly, weekly or daily restrictions.   Decision 83 did not mention co-pilots, but since 1935 it has been accepted to include them and today they are covered by the same hours regulations as pilots.

As a war measure, the ALPA agreed to an extension of the domestic hours limitation to 100 hours per month, and made concessions in the international field also.   A special bill was passed by Congress to that effect, which automatically terminated on June 30, 1947.[3]   The air lines used the extra hours sparingly, and did not do so at all since soon after the Japanese surrender.

1. For the record in railway labor disputes, see H. R. Northrup, "The Railway Labor Act and Railway Labor Disputes in Wartime," American Economic Review, XXXVI, June, 1946, pp. 324–343.   The other major air-line case involving the use of the emergency board procedure concerned the mechanics of Northwest Airlines, represented by the International Association of Machinists.   A board appointed *after* the strike which grounded the air line on July 4, 1946, rejected most of the union's claims, but the latter won them in later bargaining anyway.

2. The discussion of hour regulation is based mainly on Federal Co-ordinator of Transportation, op. cit., pp. 8–33; and "Employment Opportunities in Aviation Occupations," Bulletin No. 837–2, U. S. Bureau of Labor Statistics, 1946, pp. 31–32.

3. Public Law No. 535, 77th Congress, 2d Session, 1942.

The restrictions in international operations are of necessity more flexible. The Civil Air Regulations provide that a pilot may not fly more than 350 hours in any 90 days, nor over 1,000 hours in any year. When only two pilots and an additional crew member other than a steward or stewardess are carried, pilot flight time is limited to 120 hours in any 30 days, 300 hours in any 90 days, and 1,000 hours in any year. Restrictions are even more rigid if the technical crew consists of only one or two pilots.

In both domestic and international service, actual flying time of pilots averages closer to 80 than to 85 hours, since in order to remain within the limitations, pilots must often be scheduled less than the maximum. Efforts of the ALPA to reduce the 85-hour limitation by collective bargaining have thus far been unsuccessful.

Besides monthly limitations, the Civil Air Regulations also limit daily and weekly flying hours. Thus, a pilot in domestic operations is not permitted to fly more than eight hours during any consecutive twenty-four without a rest period. If a pilot is forced to exceed this limit because of bad weather or crowded airport conditions, he must be given a twenty-four hour rest period before assignment to any flight or ground duty. Domestic pilots are, moreover, limited to thirty flight hours per week, and must have one day of rest in seven. In international operations, rest periods are likewise required and limits are set on daily flying hours. These requirements are, however, quite flexible, especially when, as is typical, more than two pilots are carried.

The restrictions on flying time are apparently applicable to all common air carriers, whether certified or not. Enforcement is just commencing now in the uncertificated field, with wide variation of practices reported.[4]

Besides flying time, pilot crews have to spend some pre-flight preparation and briefing time and post-flight checking-in time on the ground. In addition, they are at times grounded by inclement weather or mechanical defects. ALPA officials have estimated that total ground time averages as high as an hour for every hour in the air, but the records of the carriers show an average of twenty to thirty minutes ground time per one hour of flight time.[5]

4. U. S. Bureau of Labor Statistics, op. cit., p. 36; American Aviation, September 15, 1946, p. 20.

5. TWA Emergency Board, Carriers Exhibits Nos. 36, 120, 136, and 137; U. S. Bureau of Labor Statistics, op cit., p. 32.

Besides restrictions on flying hours, the ALPA has repeatedly sought mileage limitations, basing its argument on alleged increased fatigue resulting from increased mileage flown, and on the need to prevent technological unemployment. Both the NRA National Labor Board in Decision 83 and the more recent TWA emergency and arbitration boards rejected ALPA proposals of this type as unwarranted.[6]

As for technological unemployment, there is as yet no evidence of its existence for pilots in the rapidly expanding air transport industry. In 1933 average air-line aircraft speed was approximately 120 miles per hour, and total pilot and co-pilot employment approximately 500. In 1945 the comparative figures were approximately 170 miles per hour and 5000.[7] Since 1945, both average speed and total employment have continued to increase. In November, 1946, more than 4,000 pilots and co-pilots were employed on four-engine equipment alone.[8]

The ALPA argues, in effect, that if bigger, faster planes had not been introduced, more pilots would be employed.[9] This assumes that the public would patronize the air lines to the same extent that it does today if 1933 model planes were still in service. Actually, as demonstrated, improved equipment has created jobs for pilots, both by making flying a more attractive method of transportation and, especially in the international sphere, by extending air transport into new areas. Moreover, there is considerable evidence that the demand for air transport is elastic.[1] If that is correct, larger planes, by reducing pay load costs and permitting further passenger and freight fare reductions, will create still more pilot jobs.

RULES AND WORKING CONDITIONS

In addition to wage and hour provisions, the ALPA agreements provide for a variety of rules and benefits which govern pilot-carrier relations and add to the well-being of pilot crews. Among the more important are seniority, moving and transfer

6. TWA Emergency Board, Report, pp. 32–33; Arbitration Board, p. 22.
7. CAB Statistics, diagramed in TWA Emergency Board, Carriers Exhibit No. 63.
8. Reports to Airline Negotiating Conference.
9. See, e.g. the statement of President Behncke, TWA Emergency Board, Vol. I, pp. 306 ff.
1. J. L. Nicholson, "Possibilities for Lower Airline Costs," Law and Contemporary Problems, Winter–Spring, 1946, pp. 452–58.

expenses, foreign station allowance, missing and interned allowance, expenses away from base, vacations, sickness and injury pay and leave, pensions and insurance. Many of them were issues in the 1945–1947 four-engine dispute.

Seniority. Promotions and demotions of pilots and co-pilots are governed by seniority, the senior man having the contractual right to bid on preferred jobs. A pilot transferred to non-flying or supervisory duty retains and accrues his pilot seniority as long as he maintains his pilot competency certificate — i. e. as long as he is qualified to act as a commercial pilot. Temporary incapacity or loss of certificate does not detract from accrued seniority.[2]

Vacancies for first pilots and reserve pilots are bulletined and must be filled with the senior "sufficiently qualified" pilot. If no bids are made, the company may fill pilot bids from the ranks of the reserve pilots; if there are no qualified reserve pilots available, the company may assign the job to the "most junior first pilot on the System Seniority List."

Co-pilots' job vacancies are generally not bulletined, but are filled by the company. Co-pilots are promoted according to seniority, to reserve, and first pilots, but must qualify for promotion. Failure to qualify may mean dismissal. No co-pilot possesses seniority rights during his first year. Moreover, co-pilots furloughed because of lack of work accrue no seniority during their layoff, and lose the seniority accrued prior to the furlough, if its duration extends beyond one year.

In case of a merger of two air lines, the ALPA has usually demanded that the seniority lists be dovetailed, but on at least one occasion it requested only fifty per cent seniority credit for pilots of a small air line which was absorbed by a larger one.[3] It has opposed air-line mergers before the CAB, which must pass on them, until satisfactory seniority arrangements were concluded.[4]

Moving and Transfer Expenses: All ALPA agreements provide that the carrier shall pay moving expenses, including a travel allowance, living expenses enroute, and the cost of shipment of household goods for pilots or co-pilots who are forced to change their domicile at the request of the company. In addition, nearly

2. This section is based mainly on the various agreements between ALPA and the scheduled carriers, supplemented by interviews with carrier and union officials.
3. See Puffer, Air Transportation, pp. 567–58, for the case in point.
4. See, e.g. Air Line Pilot, February, 1940; February, 1941; July, 1944.

all agreements provide for similar arrangements for a change resulting from a successful pilot bid for newly established runs or bases. The amount of expenses paid is sometimes limited, but the limits are quite generous (e.g. in many contracts for shipment of household goods, actual expenses are paid up to 800 cubic feet or equivalent).

Foreign Station Allowances. The agreements covering international operations make provision for special allowances in case the cost of living in foreign domiciles is unreasonably in excess of that in continental United States. In addition, some agreements provide for allowances to cover excess currency depreciation in foreign countries.

Expenses Away From Base. All the ALPA-carrier agreements provide lodging, meals, and transportation for pilots while away from their base, whether in the regular course of their route or because they are forced to layover as a result of inclement weather conditions or mechanical failures. Most agreements permit "reasonable and necessary expenses" for these purposes, but some stipulate the actual amount allowable for various items such as meals. In international service, some contracts provide free laundry expenses and others valet services, "when circumstances require it."

Missing and Interned Allowance. If pilot crews in international service become missing or are held prisoner by hostile tribes or warring nations, their beneficiaries have generally been entitled to monthly payments equal to two-thirds to three-fourths of their previous six-months average monthly earnings for a minimum period of two years, if their whereabouts continue unknown, until they are proved dead or until they are returned to duty. The TWA arbitration board liberalized the amount to full pay equal to the pilot's last six months average monthly earnings.

Vacations. All pilot agreements provide for vacations with pay. For first pilots in domestic operations, the vacation is generally two weeks after one year of service, and two weeks annually thereafter. Vacation pay varies from base pay to base pay plus a specific allowance, which may exceed the pilots' average earnings. The length of vacation of co-pilots is determined in the same manner as that of first pilots, and they generally receive their regular salary as vacation pay.

In international service, pilots generally receive one month's vacation at their regular salary. The Pan American-Grace agreement, however, provides for sixty days' vacation after three years of service. In order that pilots in international service may enjoy their vacation at home, vacation for pilots domiciled outside the country is often deemed to begin only after the pilot has been transported to a port within the United States.

Sickness and Injury Leave and Pay. Sick leave has been paid by most companies as a matter of policy, and is now written into the Pan American and United contracts. The TWA arbitration board awarded complete sick and injury care to pilot crews incapacitated in line of company duty abroad.

The ALPA has pressed for the inclusion in contracts of a provision requiring the payment of workmen's compensation benefits to injured pilot crews, regardless of the legal coverage of such legislation. Thus, Section 28 of its agreement with Pan American provides for compensation payments by the company "in amounts not less than those prescribed by the Longshoremen's and Harbor Worker's Compensation Act, as amended, or the Workmen's Compensation Law of the state having jurisdiction, whichever act provides the higher benefits."

Pensions and Insurance. Benefit plans are not common in the air transport industry. A few air lines have inaugurated contributory pension schemes, but they are by no means general, and have not usually been incorporated in collective agreements. Company-paid life insurance is also uncommon, although both Pan American and Pan American-Grace have well-developed programs. The agreements of both these concerns with ALPA require some notice to the affected pilots before any change is made in the insurance program, and the Pan American agreement expressly grants ALPA the right to negotiate if the company changes the amount.

During the war, pilots on military contract operations were excluded from the coverage of private life insurance. Under these circumstances, army and navy contracts with air carriers provided for life insurance protection to compensate for the temporary invalidation of private insurance. The TWA arbitration award requires the company to reimburse pilot crews for any increases in premiums resulting from their transfer to international operations, but at present no extra premium for international work is charged.

Grievance Machinery. All ALPA-carrier contracts provide for elaborate grievance machinery. Grievances are taken up by the union and/or pilot and designated company officials, from whom an appeal may be made to a system bi-partisan adjustment board, composed of two carrier and two ALPA representatives, which is established by the contracts. The Pan American agreement stipulates that in the event of an adjustment board deadlock, the board shall, within thirty days, appoint a neutral referee, or if it cannot agree on an appointment, request the National Mediation Board to select the neutral member. The TWA arbitration board ordered a similar procedure. All other ALPA agreements provide merely that in case an adjustment board deadlocks, it shall have no further jurisdiction of the issue. In actual fact, however, adjustment boards under these circumstances have frequently agreed to select referees, or to have the National Mediation Board do so.

The grievance machinery which has been developed in the industry is modeled on that which existed in the railway industry prior to the establishment of the National Railroad Adjustment Board in 1934. The 1936 amendments to the Railway Labor Act provided for the establishment of a National Air Transport Adjustment Board, if and when the National Mediation Board deemed it appropriate. There has not been any serious pressure from carrier or employee groups for such an agency. The former are fearful that it would develop policies and procedures based on the much-criticized National Railroad Adjustment Board.[5] The air transport unions do not feel that the agency as proposed makes adequate provision for representation of various employee classes. The act calls for only a four-man board, composed of two carrier, one pilot, and one ground crew representative. That would not only exclude numerous non-pilot groups, but it would compel the ALPA to share authority with another organization on equal terms, a suggestion which it does not contemplate favorably. So long as the present system adjustment boards continue to function with reasonable satisfaction, there appears little likelihood that a National Air Transport Adjustment Board will be established.[6]

5. See, e.g. W. H. Spencer, The National Railroad Adjustment Board, Chicago, 1938; and Administrative Procedure in Government Agencies, Senate Doc. 10, Pt. 4, 77th Congress, 1st Session, 1941.

6. Most union-carrier agreements in the industry covering non-pilot personnel have established adjustment boards under rules similar to those in ALPA-Pan American agreement.

THEORETICAL BASES OF HIGH PILOT PAY

The ALPA bases its case for high pilot pay principally on the following factors: qualifications and skill of pilots; short pilot working life; hazards; and pilot share in technological progress. This section examines the validity of the arguments in the light of the available facts.

Qualifications and Skills. Piloting an airplane is an exacting job, which involves great responsibility for life and property and the ability to think and act with speed and courage in emergencies. "Pilots are not supermen, but they must be well endowed."[7] Professor Ross A. McFarland of Harvard University, the outstanding authority on physiological and psychological effects of flying, testified that piloting "does not necessarily demand a person of unusual physical make-up, but only normal men who are well endowed mentally and physically with emotional traits of stability and poise."[8]

Minimum requirements for pilots, as established by the Civil Aeronautics Administration, require detailed aeronautical knowledge and skill, plus at least a restricted radio-telephone operator's permit from the Federal Communications Commission.[9] First Pilots must possess a CAA transport pilot rating, and flight instructors a special instructor rating, both demanding more expert knowledge. All ratings must be renewed every two years, and transport pilots must take a physical examination every six months. A pilot who falls below minimum physical standards loses his certificate, unless "his aeronautical experience, ability, and judgment compensate for his physical deficiency."[1]

At the present time, the scheduled air lines have set standards which are considerably above the maximum legal requirements, and will undoubtedly continue to do so as long as the supply of trained pilots exceeds the demand. Today most air lines require applicants to have at least 2000 flying hours on multi-engine aircraft, to be 21 to 32 years of age, between 5 feet 7 inches and 6 feet 2 inches in height, and within a weight range of 140 to 200 pounds. In addition, some carriers require two years of college, instead of the legal minimum of high school education or its

7. TWA Emergency Board, Report, p. 11.
8. Ibid.
9. This may soon be replaced by a CAA radio-operating authorization.
1. Civil Air Regulations, Part 29-2, quoted in U. S. Bureau of Labor Statistics, op. cit., p. 10.

equivalent.[2]   Much stress is also placed on general appearance, with the "clean-cut Anglo-American" type generally preferred.

The cost of obtaining experience enough to qualify as a candidate for a scheduled air-line pilot position may be considerable or it may involve little or no cost at all. It is considerable if the experience is obtained out of one's own pocketbook without recourse to the Army or Navy schools or without being obtained in a non-transport piloting position.   On the other hand, the bulk of present-day pilots and presumably most of those for some time in the future, are graduates of Army and Navy flying schools, and hence not only obtained their training without cost but were paid a salary while receiving it.

Exacting as the requirements for piloting positions are, they fall far below those for most professions in all except the physical qualifications, and even the poise and emotional balance which pilots must possess cannot be lacking in many professions, notably the medical, legal, and teaching ones.   Yet average pilot pay exceeds that in all but a few professions.[3]   Moreover, the cost of acquiring proficiency in most of the professions is considerably above the expenses which the average pilot incurs to learn his trade.

The positions most comparable with those of airplane pilots are probably locomotive engineer and steamship captain.   In 1945 local and way-freight engineers (who had the highest average of an engineer's group) had average annual earnings of $5,479; air-line pilots employed on two-engine equipment in domestic service averaged $8,625 for the same period.   In 1945 ship captains employed on vessels owned or under charter to the War Shipping Administration received on ships of 17,000 gross tonnage (the highest-paid classification, except for masters of super liners) a rate of $505 per month.   Air-line pilot pay for a first-year pilot on

2. U. S. Bureau of Labor Statistics, op. cit., p. 12.

3. The November, 1944, Medical Economics Magazine reported average physician net income in 1943 as $8,688, the net figure having been arrived at by deducting certain basic expenses which would put gross two to four thousand higher.   According to a U. S. Bureau of Labor Statistics survey summarized in the Monthly Labor Review, June, 1946, pp. 879–884, the 1943 median annual earnings of chemical engineers with 36–40 years' experience was only $6,620.   It is not necessary to adduce data to show that college professors in 1945 did not average $8,625, the average annual earnings of domestic pilots in that year.   Pilots, it should be noted, have approximately the same amount of leisure as have college professors who teach the most reasonable schedules.

two-engine equipment flying 85 hours per month, half day and half night, was $673.08.[4]

Within the air-transport industry, pilot pay compares very favorably with that of other highly skilled or professional personnel. In 1945, 59.7 per cent of the pilots employed by twelve air lines averaged $700 per month or more, whereas only 1.0 per cent of the aeronautical engineers, no meteorologists, and 18.2 per cent of the company physicians averaged as much or more.[5] That air-line pilots are very well compensated by any reasonable standard for their skill and qualification is evident.

Length of Earning Life. The ALPA has long made much of the fact that the earning life of pilots is short and hence pilots must be compensated for the fact that they cannot expect many active years in the profession. Actually, however, the working life of the average pilot may now be expected to last until he attains his fiftieth birthday. Again, to quote Dr. McFarland:

> After an air-line pilot has flown for many thousands of hours, he possesses training and judgment in the air which cannot be purchased at any price. Any air line would give large sums to prolong the useful lives of the older pilots. One of the largest operating companies has over one hundred pilots who are over 45 years of age. Most of these men are in excellent health and are doing their best work at 45 to 50 years of age. It is recognized that they are a highly selected group and that if they live wisely with regard to diet and exercise, and the excessive use of alcohol and tobacco, they may be able to fly until they are 50 to 55 years of age.[6]

Statistics of average pilot age are not very meaningful, in view of the youthful character of the industry. For example, statistics presented to the TWA Emergency Board showed that ages of 222 pilots employed by TWA averaged 33 years and 3 months on December 31, 1945; but the averages ages of 62 meteorologists and 2,854 mechanics employed by the same air line in May, 1946, were 28 years and 10 months, and 31 years, respectively.[7] We must conclude that the working life of a pilot is likely

4. TWA Emergency Board, Carriers Exhibits No. 16 and 52.
5. TWA Emergency Board, Carriers Exhibit No. 44; see also U. S. Bureau of Labor Statistics, op. cit., pp. 17–19, 35, for comparisons of pilots with dispatchers and meteorologists.
6. Ross A. McFarland, "The Older Worker in Industry," *Harvard Business Review*, Summer, 1943, p. 514.
7. TWA Emergency Board, Carriers Exhibits Nos. 57–61. The TWA statistics are typical of those of other air lines. As of December 31, 1945, the average age of 3,770 pilots and co-pilots employed by the thirteen air lines involved in the case was 31 years and 10 months.

to be longer than is popularly supposed, but ten to fifteen years shorter than most occupations or professions. An offsetting factor, however, is the general air-lines practice of selecting management personnel from pilots. Such promotion to better paying executive jobs will undoubtedly occur frequently as long as the industry continues to expand.

Hazards. The ALPA has laid great stress in wage negotiations on the argument that piloting is a hazardous occupation, and Decision 83 was adopted to a large extent on that basis. In 1932 the rate of fatalities per thousand among pilots and co-pilots was 35. That rate has steadily decreased, and in 1945 was only 2.6 fatalities per thousand. The lower fatality rate is reflected in pilots' insurance, rates which in 1931 were $50 extra premium per thousand, whereas the extra premium in 1945 and 1946 was only $3.00 and $2.50 per thousand, respectively.[8]

The hazards of pilots compare favorably with other similar employments; for example, the fatality rates for 1945 for railway road passenger engineers was 3.2 per thousand; for yard brakemen and helpers, 3.4 per thousand. Similarly, extra premiums for insurance in these railroad occupations and in numerous other occupations and industries exceed those which pilots are now forced to pay.[9]

The ALPA has claimed that the new and larger four-engine equipment increases pilot hazards. The difficulties with the Constellation engine which resulted in this type being temporarily grounded by order of the Civil Aeronautics Administration appeared to support the ALPA claim. These defects, however, have been fully corrected, according to the Civil Aeronautics Board.[1] Moreover, the Naval Air Transport Service, in nearly three years of operation, found the DC–4 was involved in only one-tenth as many fatalities as the DC–3.[2] The fact that four-engine

8. TWA Emergency Board, Carriers Exhibits Nos. 27, 28, 75 and 76, Statistics from Civil Aeronautics Board and Connecticut Life Insurance Company.

9. TWA Emergency Board, Carriers Exhibits Nos. 29–31. Recently the New York Life Insurance Company announced that $2.50 extra premium would apply only to the first ten years of an insurance contract, with no extra charge thereafter.

1. New York Times, February 13, 1947.

2. TWA Emergency Board, Carriers Exhibit No. 32. During this period, the NATS flew DC–4's 65,763,341 transport miles with only one fatality; and DC–3's 89,398,793 transport miles with fourteen fatalities.

planes can climb above much bad weather and are equipped with the latest safety devices would appear to support a conclusion opposite to that reached by the ALPA. Such a conclusion is borne out by the most recent accident data. In 1946, the Civil Aeronautics Administration announced that the fatality ratio for passengers per 100,000,000 passenger miles flown on domestic air lines was only 1.2 as compared with 4.8 in 1935 and 2.1 in 1945.[3]

On the other hand, airport facilities have not been completely equipped to withstand the tremendous postwar expansion of air travel, and modern safety devices, e.g., radar, have not been utilized as fast as they might be. Moreover, it cannot be denied that the pilot bears a tremendous responsibility for life and property, which is reflected in tensions or mental hazards not measured by fatality rates. To what extent these difficulties and tensions are great enough to overbalance the lowered fatality rates as a theoretical basis for high pilot earnings is, of course, not subject to measurement.

Technological Advancement. Much of the ALPA argument before various wage dispute boards has been devoted to the question of pilots' share in advancing technology, based on an analysis of the "pay load per ton mile." Thus ALPA witnesses devoted considerable energy before the TWA Emergency Board to demonstrating that the new four-engine equipment carries a larger payload at less cost than does the DC–3, and therefore that pilots should receive a share of the savings resulting from increased "productivity" of the new equipment. ALPA also advanced as a corollary argument the claim that flight miles should be limited in order to cushion the unemployment which they prophesied would result from the more productive equipment.[4]

Decision 83, by automatically providing for increased rates for faster planes, and the TWA Emergency Board by implementing the mileage and hourly pay components of Decision 83, provided for pilot sharing in increased productivity. Whether pilots should receive a larger share, as the ALPA demands is open to debate. There is, however, evidence that the pilot share is generous. In December, 1944, pilots were 3.76 per cent of all

3. New York *Times*, January 29, 1947.
4. TWA Emergency Board, Testimony, pp. 482 ff.; Pilot Exhibits Nos. 10–14.

air-line employees, and received 14.38 per cent of all wage payments in the industry.[5]

As for technological unemployment, we have already noted that far from causing unemployment, larger, faster planes have created jobs for pilots by making air transport a more attractive method of travel, by opening up previously inaccessible areas to the industry, and by creating possibilities for further fare reductions.

Paradoxically, by attempting to prevent "unemployment" where employment is being created, the ALPA may find itself the cause instead of the cure. Prior to 1945, pilot wages did not loom important in managerial decisions regarding the use of new equipment. But the heavy wage demands and the militant tactics accompanying the 1945–1947 four-engine dispute have had important repercussions. As a result of the pilot strike in October-November, 1946, TWA cancelled large orders for new planes and laid off the pilots who had been employed to operate them, and has been all but driven into bankruptcy. The demand for pilots is certainly less inelastic than formerly, and the ALPA could make it elastic. When one notes that ALPA wage demands are certain to be reflected in the demands of the other air-line employees, who are now organizing rapidly, that possibility is materially enhanced.

Bargaining Power. Probably the most meaningful explanation of pilot earnings is found in the bargaining power of the ALPA. A very strategically situated group, it has used its advantage to the utmost. In the face of an extraordinary large surplus of trained pilots which the war created, it has continued to boost the salaries of its members. Moreover, it has been unusually adroit in the political field. Despite its limited numbers (and votes), it has won the aid of Congress to a degree which has been exceeded by few organizations. Taking full advantage of the "romantic allure" of the industry and jobs, pilots, lobbying in their smart uniforms, have impressed the legislators time and again. The latter's sympathies have been obtained more easily because of the fact that they were also aiding the carriers with public funds. The results — Decision 83 written into law — are impressive indeed. Whether it will have similar success in the future remains to be seen. Certainly, the TWA strike was not

5. TWA Emergency Board, Carrier Exhibit No. 42, based on data compiled by Civil Aeronautics Board.

helpful. For, in the eyes of the public and the Congress, it reduced a group considered "professional" to the regular employee level and stripped them of their glamour. To what extent this will reduce their influence to the number of votes they can cast remains to be seen.

Conclusion

Despite the high salaries and good working conditions on the air lines, industrial relations at the present time can scarcely be termed satisfactory from the point of view of either party or the public. Within the last two years negotiations between the pilots' union and the various air lines have been conducted on a marathon basis and have been interrupted by one serious labor stoppage. Future prospects will, of course, depend upon many circumstances. There are important issues upon which pilots and carriers are in firm agreement. They include the continued improvement of safety conditions, the maintenance of private ownership as against government ownership utilizing armed service or civil service personnel, bettering competitive position of the air lines within the general transportation field, and the continued use of former pilots to fill executive positions.

On the other hand, there are factors which may tend to worsen the possibilities of industrial peace. Decision 83 itself may be one of these. It provides for automatic increases under the highest minimum wage law in the land; but like any other sliding scale, it does not give proper political acknowledgement to union leaders for securing these increases, and one is impressed by the fact that ALPA leadership feels the need of acknowledgement of that character. The insistence of the union on long-drawn-out negotiations and a new bargain every time a new plane is introduced and on squeezing the industry as hard as it can are certainly factors which make good industrial relations difficult. The bad feeling engendered by these tactics in the TWA four-engine case may survive for some years to come. Of particular importance in this connection will be the question whether the ALPA continues to attempt to deprive the carriers of the right for which the ALPA so vigorously, successfully, and rightfully fought — namely, the right to choose the bargaining agent without interference from the other party.

Within the union there are other factors which may upset the

present equilibrium. These include opposition to Mr. Behncke's leadership and dissatisfaction among co-pilots with their lack of effective voice in union affairs. The extent to which either may assert itself remains to be seen. Finally, it should be noted that industrial relations in the air lines cannot be disassociated from the general industrial relations picture. The four-engine dispute was conducted when strikes were in the air. Today labor peace is the prevalent tone. It permeated the airlines when no difficulties arose over the introduction of the new DC-6. If the pilots and carriers can continue to settle their problems before the calm atmosphere dissipates itself, perhaps they can restore mutual respect and a common will to peace.

HERBERT R. NORTHRUP.

NEW YORK SCHOOL OF SOCIAL WORK,
    COLUMBIA UNIVERSITY

SMALL-SCALE INDUSTRY IN JAPAN

SUMMARY

I. The rôle of small-scale industry: the prewar situation, 578; productivity, 582; trends, 584; war-time changes, 586; the social context, 592. — II. Prospects for large-scale industry: effect of air raids, 593; occupation policies, 593; other unfavorable factors, 594; the textile industry, 596. — III. Prospects for small-scale industry: advantages of small-scale plants, 597; obstacles to survival, 599. — IV. Implications: absorption of labor, 601; development of a strong middle class, 602; needed reforms, 602; the labor movement, 603; occupation policies, 603.

In the process of reshaping the Japanese economy to peaceful and democratic ends, the Allied Occupation is influencing the character of virtually every economic institution, sometimes by direct measures and sometimes by the unintended consequences of other policies. One of the developments which has not yet received much attention is the altered rôle of small-scale industry.

Before the war, industrial production in small shops, especially in tiny domestic workshops, was a notable feature of the Japanese economy. While these small shops resembled their counterparts in other countries with respect to many of the technical methods and operating problems, they acquired a special character in Japan because of their widespread extent, their persistence, and their conservative influence on the social structure. Unlike the situation in the United States, where "small business" is often championed as a bulwark of democracy, small-scale industry in Japan lacked the attributes which have gained support for the American institution.

Under the economic circumstances of the post-Surrender period and the policies of the Occupation authorities, small-scale industry is likely to expand. This article assesses the probable rôle of small-scale industry against the background of its earlier history. Section I is concerned with the prewar and war-time trends, introducing new data drawn from Japanese sources and from the findings of the United States Strategic Bombing Survey, of which the writer was a member. The other sections deal with the current situation, covering the prospects and significance of an expansion of small-scale industry in the new Japan.

I. The Rôle of Small-Scale Industry

For the purposes of this discussion, small-scale industry is defined as manufacturing and processing establishments with less than fifty workers.[1] Within this group in past years a certain variety of types has been observed. Some shops had a fair amount of power-driven machinery, while others depended on handicraft. Some were conducted in homes with family labor, including the part-time work of school children and housewives, while others occupied small factory buildings and used hired help. Some were operated on a regular basis all year round, but others were intermittent or seasonal activities complementary to other occupations such as agriculture or fishing. Some turned out finished products for sale to ultimate consumers, while others specialized in components which were supplied to outside factories on a sub-contracting basis. Some produced traditional goods for local consumption — for example, floor mats (*tatami*), wooden clogs (*geta*), and various foodstuffs; others manufactured Western-type items for export as well as domestic sale — for example, lacquer wares, pottery, light bulbs and bicycles. Some of the small shops were independently financed and operated, but others were linked with large business organizations or were financed by commercial capitalists. Thus small-scale industry ran the gamut from the most primitive kind of handicraft to small mechanized factories; and from independent business units to cottage industries which in effect did outwork on a piece-rate basis.[2]

Despite such variations, the small-scale industries were clearly distinguishable — in type as well as in size — from medium and

1. It is noteworthy that the term "small-scale industry" refers to the size of the individual production unit, not to the size of an industry. Small-scale producing units are found even in the largest industries.

This definition of scale in terms of employment in the individual production unit is adopted in conformity with the Japanese statistics. Data are not available for determining what proportion of the producing units are integrated, horizontally or vertically, in large combines. However, in the area of small enterprises, outright integration is not favored as much as other, looser forms of organization; hence we have assumed that in the typical case the small plant is an individual proprietorship.

2. Uyeda, T., The Small Industries of Japan, Institute of Pacific Relations, Oxford, 1938; Farley, Miriam S., "Pygmy Factories — The Backbone of Japanese Industry," Far Eastern Survey, January 6, 1937; Schumpeter, E. B., et al., The Industrialization of Japan and Manchukuo, 1930–1940, New York, 1940, ch. XII–XV; Orchard, J. E., Japan's Economic Position, New York, 1930.

large-scale factories.[3] In general, small-scale shops required little plant and capital equipment, since they relied on a very high ratio of labor to the other factors of production. The mechanical equipment was simple but sometimes highly specialized and was utilized to an extent and efficiency made possible by Japan's abundant supply of electric power. Much emphasis was placed on manual dexterity, resulting in skills exceeding those of most workers in the more highly mechanized factories; at the same time, the small shops had a large proportion of quite unskilled persons (including children) to do simple preparatory or finishing tasks and heavy work. The work force, commonly built around the family as a nucleus, was held together by close personal and customary relationships, which recognized and adjusted for various interruptions to regular working. An important outcome of these traits was the fact that the small shops were capable of irregular operations, employing small and varying stocks of raw materials as well as varying amounts of labor. Finally, the products were usually sold locally— either for direct consumption or for wider distribution by the parent firm, the merchant capitalist, or the producers' guild.

These traits contrast with the characteristics of the medium and large-scale enterprises, which emphasize a higher ratio of capital equipment to labor, along with regularity of operation, finely sub-divided tasks, and distribution of the products to an extensive mass market. In this, the larger enterprises resembled their counterparts in other countries, although few became truly mass-production operations.

The special character of small-scale manufacturing units does not mean that these units were restricted to a few industries or to a small share of all industry. Table I presents for 1939[4] the distribution of employment by size of plant. Small shops as defined here represented more than half of all industrial employment, while

3. Large-scale industry in Japan is defined here as establishments employing 1000 or more operatives, while medium-scale enterprises employ 50 to 1000 operatives. This classification is based upon homogeneity of type and of trend, but is necessarily somewhat arbitrary; in certain industries the small-scale limit might more appropriately be set at 30 workers, while large-scale operations might begin with 500 employees.

4. The year 1939 was selected for analysis because in that year, for the first time, the Japanese statistics covered all sizes of plants; and because subsequent years were affected more and more by the special distortions of the war effort. The record before 1939 and in subsequent years is discussed below.

TABLE I

PERCENTAGE DISTRIBUTION OF INDUSTRIAL EMPLOYMENT BY SIZE OF PLANT AND BY INDUSTRY, 1939

	Size of Plant, by Number of Workers Employed				Relative Employment In Each Industry
	Total	1 – 49	50 – 999	1000 and Over	
All Industries	100.0	52.6	27.8	19.6	100.0
Textiles	100.0	44.2	42.4	13.4	23.1
Metals	100.0	41.5	27.4	31.1	10.6
Machinery and Tools	100.0	25.5	34.7	39.8	25.4
Ceramics	100.0	65.8	29.8	4.4	3.8
Chemicals	100.0	35.9	33.8	30.3	9.3
Lumber and Woodworking	100.0	89.8	10.2	...	6.6
Printing and Bookbinding	100.0	73.7	22.7	3.6	1.0
Food Processing	100.0	83.8	14.7	1.5	8.0
Gas and Electricity	100.0	56.2	43.8	...	.3
Miscellaneous	100.0	86.2	12.3	1.5	11.9

Source: Industrial Statistics (Kogyo Tokeihyo), Ministry of Commerce and Industry, Tokyo.

medium-sized enterprises claimed little more than one-fourth and large-scale less than one-fifth. Small-scale manufacturing was found not only in the manufacture of traditional Japanese-type goods and handicraft items, but also in industries which in Western countries tend toward a larger scale of operation. As Table I indicates, small shops in 1939 constituted 44 per cent of the textile industry,[5] 42 per cent of the metal-working, 36 per cent of the chemicals.[6] Of course, the percentage was highest in the peculiarly traditional or handicraft trades, such as ceramics (66), lumber and woodworking (90), food-processing (84) and miscellaneous trades (86); but only in the manufacture of machinery and tools did the percentage (26) fall far short of the all-industry average of 53 per cent.

Within the broad categories, the small shops tended to concentrate in certain branches. Thus in textiles the small-scale branches were throwing, weaving, hosiery-knitting, dyeing, bleaching and finishing, and a few other operations, which together provided about one-half of all employment in the textile industry; while spinning and cotton broad-woven staples were in the hands of a few large firms. Similarly, in metal-working the small-scale branches were casting, miscellaneous metallic goods, and plating — branches which accounted for more than half the whole metal-working category; but iron and steel, copper-refining, etc. were large-scale operations. In machinery and tools, small shops were chiefly devoted to the manufacture of the simpler implements and components of complex machines. In chemicals, they were important in paints and varnish, oil and tallow, celluloid, leather prepa-

5. In the Japanese classification, "Textiles" refers to spinning and weaving and associated activities. The manufacture of clothing is included in the "Miscellaneous" category.

6. For a suggestion of the order of magnitude of these percentages, the following figures on United States industry in 1939 are presented:

Industry	Employment in plants with less than 51 workers as a per cent of employment in all plants in each industry
Textiles (excluding apparel)	5.2
Iron and Steel and their products	9 7
Non-ferrous metals and their products	21.2
Chemicals	27.9
All manufacturing industries	20.7

These percentages are calculated from the Sixteenth Census of the United States, Manufactures, Vol. I, using the data on wage-earners, adjusted for entrepreneurs who are operating workers. The United States figures should not be compared directly with the Japanese statistics, because of differences in definition and classification.

ration and other light chemicals production, aggregating about one-fourth of all chemical employment.[7]

It is worth noting, in contrast, how the medium and large-scale enterprises were distributed among the various industries. Medium-sized factories, which in 1939 claimed nearly 28 percent of all industrial employment, constituted more than this average percentage in textiles, machinery, chemicals and public utilities. In the case of large factories, which accounted for nearly 20 per cent of all industrial employment, emphasis was concentrated on metals, machinery and chemicals; in textiles, less than 14 per cent, and in other industries practically negligible shares of employment, were claimed by the large units.

Attention to the statistical breakdown of plants by number of employees should not obscure the functional relations between the various size-classes. While in such industries as food-processing the small plants often could operate quite independently, in other industries, such as metal-working, the little shops were obviously dependent upon the larger establishments for their raw material as well as for part at least of their equipment, markets and financing. In cotton textiles, the spinning of the yarn was concentrated before the war almost completely in the great mills of twenty-three firms (which have now been reduced to ten firms as the result of war-time amalgamations); but at the weaving stage, the yarn was used by numerous small independent units.[8] In the manufacture of electrical equipment, a high degree of specialization was attained by small plants, which fed components into master factories.

Another vital consideration is the productivity of the small plants, i.e. their ability to contribute to the goods and services available for consumption, capital formation, the balance of foreign payments, etc. Little enough is known about this subject, even for the most advanced industrial nations; in the case of Japan, only the most rudimentary analyses have been made. The usual finding has been that Japanese small-scale shops rank lower in value of output than in number of workers, and the conclusion has some-

7. Data for the above are drawn from the Industrial Statistics (Kogyo Tokeihyo), 1939. The concentration of the small shops in certain lines within the broad categories is important for the question of the ability of the little units to stand alone; see discussion in Section II.

8. The Textile Mission to Japan, Department of State, Publication 2619, Far Eastern Series 13, 1946.

times been drawn that small-scale industry is less efficient than larger establishments.[9]

This method of analysis is quite dubious. The experience of other countries has demonstrated that value-of-output figures are a very inadequate measure of efficiency.[1] In fact, alternative methods, such as analysis of costs of production for identical products, and analysis of rates of return on invested capital for profitable businesses only, have demonstrated that efficiency is not directly correlated with increasing size of plant.[2] There is an important distinction to be drawn here between technologic efficiency and commercial success, although the two are frequently confused. This distinction is particularly relevant to Japan, where the commercial environment has been decidedly adverse to the small businessman. All observers agree that the entrepreneur and his workers were able to extract only a very low standard of living from the operations of small shops. In part this performance may be attributed to lack of managerial skill, but still more importance may be attached to other factors: the ability of money-lenders, landlords, suppliers of materials and parent factories to exploit the small firm, the cut-throat competition among the multitudes of little units, the burden of severe taxation, the direction of the Government's financial and other aid in the past primarily in favor of large companies, and the barriers to modernization resulting from difficulties in accumulating or borrowing funds.

In the technologic operation alone, the Japanese have shown themselves very ingenious at adapting their system of small-scale production to the lines in which they have special advantages. In the case of Japanese-type goods, whose production is overwhelmingly small-scale, it appears that large factories were often technically inferior; because of consumer specifications and the require-

9. See, for example, Farley, op. cit., p. 6, and Utley, F., *Japan's Feet of Clay*, p. 161.

1. Under "General Explanations," the Census of Manufactures for the United States, 1939, points out that the value-of-output approach is vitiated by failing to take account of the value of the materials and products entering into each successive stage of the productive process (p. 5). There are other limitations, particularly the influence of market valuations of diverse products, which further reduce the usefulness of value-of-output figures for comparisons of efficiency by size of plant; and these limitations also apply to "value-added" figures, which Japan does not have.

2. See TNEC Monograph No. 13, *Relative Efficiency of Large, Medium-sized and Small Business*, and Crum, W. L., *Corporate Size and Earning Power*.

ment of high quality, together with the handicraft character of the processing and its ready divisibility into separate stages, small shops were the appropriate unit. In the case of large-scale assembling operations, small shops developed a high degree of special skill in producing component parts for the parent plant. In still other lines, many a small shop found that the operation of a single machine-unit could be at least as efficient as the operation of a long line of such units, if there was no concatenation between them. In general, it appears to be true that neither technologic efficiency nor commercial success is a simple function of size in the Japanese economy any more than in other countries. The prospects for small-scale industry in the postwar period will be dependent in large part upon the kind of commercial conditions which develop under the Allied occupation.

Before turning to these prospects, it will be useful to review the trends in small-scale industry in Japan, starting with 1930 and carrying the record through the war. The extent of small-scale industry in 1930, great as it was, seems to have been exaggerated by most students of the Japanese economy. Because of the lack of statistical data on the smallest shops (those employing fewer than five operatives), it was necessary to rely on estimates by Japanese statisticians. As a result, it has been generally assumed that in 1930 small-scale shops accounted for almost three-fourths of all industrial employment, the 1–4 class alone being held responsible for over half the total.[3]   The basic data on size of enterprises came from the surveys of the Ministry of Commerce and Industry, and were published as Factory Statistics (renamed Industrial Statistics in 1939). Until 1939 these statistics were compiled from the questionnaires directed to employers who hired five or more operatives.[4] In order to secure a complete picture, Japanese statisticians made estimates for the workshops employing fewer than five workers. The usual method was a fairly crude subtraction of the 1930 Factory Statistics from the total of the industrially occupied population as enumerated in the 1930 Census of Population, the differ-

    3. E. g., Allen, G. C., in Schumpeter, op. cit., p. 622, and Farley, op. cit., p. 2, both citing Uyeda, T., The Growth of Population] and Occupational Changes in Japan, 1920–1935, Institute of Pacific Relations, p. 15.
    4. An "operative" is a person actually engaged in the productive process, whether part-time or full-time, and is distinguished from clerical, technical and supervisory employees, gatekeepers, messengers, and charwomen, etc., and from self-employed persons and family workers.

ence being taken to represent the workers in the 1–4 class.[5] Although some allowance was made for obvious differences in coverage, the two series were not carefully adjusted. It is, in fact, doubtful whether the two sets of statistics can be used together reliably.

No new direct data on the 1930 position are available, but it is now possible to show that the estimation technique probably exaggerated the extent of small-scale industry. In 1939 the Ministry of Commerce broadened the scope of Factory Statistics to include the smallest shops — those employing less than five operatives.[6] The actual count of employment in these shops, including self-employed and family workers, in 1939 and again in 1940 yielded a figure that was about half the number estimated for 1930.[7]

This large discrepancy does not of itself constitute a proof that the 1930 estimate was in error. It is possible that a great absolute and relative decline in small-scale industry occurred between 1930 and 1939. However, this supposition runs counter to general impressions of contemporary observers,[8] and is contradicted by the available statistics.[9] It seems reasonable to conclude, therefore,

5. See Asahi, I., The Economic Strength of Japan; Takahashi, K., Factors in Japan's Recent Industrial Development; Uyeda, T., The Growth of Population and Occupational Changes in Japan, p. 14; Social Reform (Sakai Seisaku Jiho), June, 1941, pp. 109–111; ibid, April, 1940, pp. 128–165; Takano, I., "Statistics in Japan and Germany in Relation to Small Size Industries," Ohara Social Research Institute Magazine, Vol 3, No. 7; Nippon Rodo Nenkan, 1938, pp. 7–9.

6. The statistics were presented in two series. One set continued to deal only with those factories employing five or more operatives, while the other set covered only shops with fewer than five operatives. In the latter group, self-employed persons and family workers were counted.

7. The number of workers in 1930 in shops employing fewer than five was estimated at nearly 2.4 million according to the Japan Labor Year Book (1937), whereas the Industrial Statistics showed only 1.2 million in 1939.

8. According to Miss Farley, writing in 1937, "Since 1930, according to the testimony of several observers, there has been a rapid increase in the number of small plants, stimulated by the wide margin of competitive advantage conferred by depreciation of the yen. It seems probable that small-scale industry has not only held its own in a period of rising employment, but has more than maintained its relative position." Op. cit., p. 2.

9. Between 1930 and 1939 every size class counted by the Factory Statistics showed a sizeable absolute increase in employment. For all factories employing five or more operatives, total employment more than doubled. (See Table II-A.) It will also be seen from Table II-A that the smallest class counted during the thirties — the 5–9 group — just about held its relative position from 1930 to 1939, and the next two classes actually improved their standing somewhat. Furthermore, as regards the 1–4 class, in the years following 1939 (the earliest date for which we have direct figures on this class),

that the earlier figures contained a considerable degree of error, and that the employment provided by the domestic shops (1–4 workers) probably was closer to the 1939 percentage which was about one-fourth of all industrial employment, than to the 1930 estimate of over half the total.[1]

The significance of this error can be seen, for example, in the opinion held by Professor Uyeda and others that Japan had as much handicraft industry in 1930 as Germany had had in 1880.[2] However, if the number of workers in domestic shops in Japan in 1930 was actually considerably smaller than has previously been supposed, the proportion of small-scale employment probably was not much larger in the Japan of 1930 than in the Germany of 1925. This does not mean that small-scale industry was not important in both countries. Both Japan and Germany present a sharp contrast to the United States, where in 1939 not quite 21 per cent of all workers were employed in establishments with less than fifty employees.[3]

The trend of small-scale industry from 1930 to 1939 must be measured in terms of factories employing 5–49 workers (Table II-A). For the years 1939–1942 data are now available covering plants in the 1–4 class along with the larger ones (Table II-B). The small plants showed a slight relative gain from 1930 to 1939, tending to level out during 1939–1941, with the first signs of a substantial decline appearing in 1942. The individual classes show

there was no strong downward trend, such as might have been expected if a decided decline had been in progress during the preceding years.

1. Some evidence as to the magnitude of the error of the 1930 estimate is obtained by applying the 1930 estimation method to 1940 data. If the method were valid, the results should not be far different from the actual count of employment in small shops. When the calculation is actually made, the results turn out to be far larger than the actual count made by the Ministry of Commerce and Industry, particularly if family workers and entrepreneurs are excluded. While it is entirely possible that there is serious under-counting of the tiny shops in the Industrial Statistics, there is a counterbalancing bias in the inclusion of family workers, many of whom worked part-time only. It therefore seems reasonable to conclude that the 1930 estimation technique yielded an exaggerated figure.

2. Professor Uyeda points out that "in 1925 only 22 per cent of the industrial workers of Germany were in plants having fewer than six employees and 45 per cent in plants having fewer than fifty"; whereas Japan was alleged to have 53 per cent of the total industrially employed in establishments having fewer than five workers, and 70 per cent in shops employing less than fifty. (Uyeda, T., The Small Industries of Japan, pp. 8–9)..

3. Computed from Census of Manufactures, 1939.

TABLE II

PERCENTAGE DISTRIBUTION OF INDUSTRIAL EMPLOYMENT BY SIZE OF PLANT

SELECTED YEARS, 1930-1942

Year	Total Employment		Size of Plant, by Number of Workers Employed							
	Number	Per Cent	1 - 4	5 - 9	10 - 29	30 - 49	50 - 199	200 - 499	500 - 999	1000 and Over
			PART A*							
1930......	1,683,563	100.0	*	11.5	15.4	7.9	21.9	14.2	11.2	17.9
1932......	1,733,511	100.0	*	12.1	16.3	8.5	21.3	13.3	11.0	17.5
1934......	2,163,453	100.0	*	11.8	16.3	8.8	20.9	12.0	10.6	18.6
1936......	2,592,687	100.0	*	10.2	17.6	9.4	19.7	12.1	9.9	21.1
1938......	3,217,715	100.0	*	10.1	16.1	8.6	17.7	11.7	9.4	23.5
1939......	3,786,247	100.0	*	11.6	18.1	8.4	16.6	10.9	8.8	25.6
			PART B†							
1939.....	4,950,881	100.0	23.5	8.9	13.8	6.4	12.8	8.3	6.7	19.6
1940......	4,986,930	100.0	22.9	9.1	13.5	6.1	12.7	8.1	7.0	20.6
1941......	4,943,319	100.0	23.4	9.1	13.7	5.8	12.4	7.8	6.3	21.5
1942......	5,061,888	100.0	22.5	7.7	12.8	5.5	11.5	8.0	6.2	25.8

Sources: 1930-1938 data from Factory Statistics (Kojo Tokeihyo); 1939-1942 data from Industrial Statistics (Kogyo Tokeihyo). Ministry of Commerce and Industry, Tokyo.

* Part A relates to employment in plants with five or more workers. Operatives only are included

† Part B relates to employment in plants of all sizes. Family workers and self-employed persons as well as operatives are included in the 1 - 4 size-class; all other size-classes include operatives only

some annual variation around this general trend, but do not depart from it widely. By contrast, the medium-sized group reveals a marked and almost uninterrupted decline: from 1930 to 1942, each of the three groups — 50-199, 200-499, and 500-999 — lost a large part of its share of all industrial employment. The most remarkable development of these years was the rapid growth of the largest-sized group, plants with 1,000 or more workers. The intense development of heavy industry and the preparation for war are responsible for the stimulation given to large-scale establishments. In a sense this stimulation was "at the expense of" the middle-sized group, but in actuality many a plant moved from the medium to the large category by expanding its facilities or increasing the number of shifts.

The small plants, on the other hand, maintained their share in the greater volume of economic activity by serving as subsidiaries to the fast-growing large factories. Even after Pearl Harbor, small industry did not vanish, and did not even decline significantly up to 1943. Thereafter a new pattern set in: units employing less than 10 operatives (these were chiefly domestic workshops) almost disappeared during 1943 and 1944 except in some traditional lines, while units employing 10 to 50 operatives actually increased their numbers in some industrial categories. The decline of the smallest-sized shops resulted from special war-time causes. The military draft frequently took the head of the household, who was usually indispensable to the domestic workshop, and deferments or exemptions even for essential productive work were rare; labor conscription often swept up the head of the household and one or more of his former assistants or employees, sending them into the large war factories. The inability to obtain materials, interruptions of transportation, scarcity of food for the ordinary civilian, and other war-time dislocations drove many household units out of business. On the other hand, the units having 10 to 50 operatives made a substantial contribution to the war effort; it was found to be both necessary and desirable to enlist their facilities as feeder plants to larger establishments, using the sub-contracting and putting-out systems. In the case of the electrical industry, this movement went so far that about two-thirds of the total output of communications equipment came from small plants.[4]

    4. The discussion of war-time developments in this and following pages is based largely on information collected by the United States Strategic

Some measure of the changes will be found in a rough statistical comparison. The 'Industrial Statistics were not collected in 1943 and 1944, but some figures for 1944 and 1945 have been published by the Ministry of Public Welfare.[5] These figures are not strictly comparable with the data presented in Table II, and the size breakdowns are also different. However, the pattern given for June 1944 is consistent with the views of Japanese government officials and businessmen that between 1942 and 1944 there was a vast increase in the relative position of the largest plants and a drastic decline in the proportion of the smallest shops, while the small shops having 10 to 50 operatives did not do badly.

During 1945, the situation changed once more, as the air attacks hit Japan.[6] The urban area attacks were extremely devastating, as has been reported many times, and the small plants were usually situated in the midst of dense urban centers. In some of the large cities which were subjected to intensive attack, destruction of small plants ran very high. In Tokyo, for example, approximately 70 per cent of the plants having less than 50 employees

Bombing Survey, including investigations made by the writer while in Japan with the Survey.

5. The percentage distribution derived from the figures for June, 1944, is as follows:

Size — Class of Factories	Percentage of Employment
Total	100.0
Less than 10 employees*	3.3
10–99 employees	20.7
100–499 employees	15.8
500–999 employees	8.9
1000 employees and over	51.3

Source: Computed from figures released by Ministry of Welfare, as reported in Jiji Press, April 3, 1946.
* Judging from the character and origination of the data, this class probably does not include the factories with less than five workers.

6. Figures on "factories destroyed during the air war," computed as a percentage of factories in operation in June, 1944, are as follows:

Size — Class of Factories	Percentage of Factories "Destroyed"
Less than 10 employees*	23
10–99 employees	21
100–499 employees	30
500–999 employees	20
1000 employees and over	7

Source: Jiji Press, April 3, 1946.
* The category "Less than 10 employees" may not include factories with less than five workers. Also, it is possible that destruction among the smallest shops was substantially undercounted.

ceased operations between October, 1944 and August, 1945, while larger plants suffered only a 20 per cent loss.[7] However, there is evidence which suggests that, for the country as a whole, the destruction of small plants was not much more extensive than the destruction of medium-sized ones — while large factories took a much lighter loss.

The largest factories were situated on the outskirts of large cities, and so escaped the full effect of the conflagrations that swept densely settled areas. In the case of the smaller-sized plants, a tour of Japanese cities soon after the Surrender revealed some repairable equipment in domestic workshops and small factories, even in areas which had been swept by fire. Furthermore, the lathes, looms and other machine tools used by farmers for off-season work, were found to be intact in most cases, since bombing did not hit the rural areas, and the equipment was often unsuitable or unavailable for removal to any large factory. In fact, some urban artisans sent their equipment to friends or relatives in the country for safekeeping, just as many large textile factories dismantled their machinery and stored it in caves.

The Ministry of Welfare report already cited contains some statistical information on employment by size of plant, for plants in actual operation in December, 1945, four months after the Surrender.[8] These statistics suggest that between June, 1944, and December, 1945, total industrial employment declined by 70 per cent; however, this change may have been due to the coverage of the surveys as well as to an actual decline in employment. Using the figures only for proportions, it appears that the share of small shops (those with less than 100 operatives in this

7. Computation based on data collected by the Urban Areas Division of the United States Strategic Bombing Survey and published in The Effect of Air Attack on Japan's Urban Economy — Summary Report, March, 1947.

8. The percentage distribution of plants in operation in December, 1945, was as follows:

Size-class of Factories	Percentage of Employment
Total	100.0
Less than 10 employees	8.2
10–99 employees	44.4
100–499 employees	24.1
500–999 employees	9.0
1000 employees and over	14.3

Source: Jiji Press, April 3, 1946.

tabulation) had greatly recovered from the preceding year, while the largest factories had dropped precipitously. However, the unreliability of the statistics in question prevents any detailed analysis.

The postwar trends may be somewhat clarified by the Labor Census of Japan taken in mid-1946. The results of this census are being published in successive sections; as the section on employment by size of firm was not available at this writing, we shall examine the report on establishments by size of firm (Table III).

TABLE III

INDUSTRIAL ESTABLISHMENTS, BY SIZE

Size-Class of Establishments	Percentage of Establishments	
	December, 1939*	June, 1946†
1 – 49 employees...............	98.7	96.3
1 – 9 employees	91.0	81.2
10 – 49 employees	7.7	15.1
50 – 999 employees	1.25	3.6
1000 employees and over	.05	.1
Total.....................	100.00	100.0
Total Number of Establishments	707,594	365,253

* Computed from Industrial Statistics, 1939, collected by the Ministry of Commerce and Industry
† Computed from Annual Labor Census of Japan, 1946, "Number of Establishments in Japan, by Industry and Size, as of June 30, 1946," Economic and Scientific Section, Research and Statistics Division, SCAP, January 2, 1947 (mimeographed). The census was taken by the Cabinet Bureau of Statistics of the Japanese Government

Comparison of the two sets of data shown above is not simple, since the figures were collected by different agencies, each using its own definitions and coverage, and operating under quite dissimilar conditions. In particular, there is reason to suspect the 1946 survey of undercounting the smallest establishments. Likewise, comparison of these data on establishments with previously discussed data on employment cannot be done readily. Finally, many of the establishments reported as existing in 1946 were in fact partly or wholly idle. Recognizing these limitations, we may note first the severe drop in the total number of establishments between 1939 and 1946, reflecting the generally low level of economic activity in Japan since the Surrender. Of this reduced total, the small shops are still the vast majority — so great a pro-

portion, in fact, that the larger establishments were able to double or triple their share of the total without seriously cutting into the predominance of the little units. The actual count in the 1946 Census showed no decline from 1939 in the absolute numbers of establishments reported in the various groups with 10 or more workers. There was an actual increase in most of these groups. Almost the whole decline in the total was drawn from the 1 – 9 group, especially the 1 – 4 class. This is to say, if we may trust the figures, that establishments with 10 or more workers increased their relative share from 1939 to 1946; the greatest gain in fact seems to have occurred among the medium-sized enterprises.

To sum up the recent history of small plants: they were not much affected by war-time changes until after 1942; as a widespread form of organization, they persisted throughout the war; individual units suffered some closings during 1943 and 1944, and suffered widespread destruction during 1945; while the end of the war undoubtedly found a great reduction in the total number of establishments (the reduction being concentrated among the smallest units), the physical and organizational potentialities for a fairly rapid revival still existed.   •

In considering the history of small-scale industry, especially with a view to future prospects, it is important to recall the social context which in the past supported this system. A major factor was the pressure of population increase against limited resources. Other prominent features were the extreme concentration of power and wealth, and the organization of society on paternalistic lines. There was practically no enterprising middle class, limited labor organization, restricted availability of capital for the growth of small enterprises to larger size, and a remarkable conservatism of traditional arts and consumption habits. In this context, small-scale shops functioned very naturally. Whether they can continue to function in an environment where some phases of the prewar situation have been swept away, while others have been intensified, is one of the questions to which the following sections are directed.

It is also important to note the contribution of small-scale organization toward the maintenance of the social system of prewar Japan. While the small shops served as sources of employment and income, they yielded — under existing conditions — only a low standard of living, and they offered little hope of social progress. Both wages and profits were lower than in larger plants, and work-

ing conditions were worse. Competition among the small shops was intense, and led to charges of dumping in foreign trade as well as exploitation of workers at home. The obstacles to unionization — multiplicity of units, small numbers of workers in each unit, paternalistic relations between employer and employees, meager margin of profits, etc. — limited the organization of this body of workers, who constituted the bulk of the industrial labor force. The income of the entrepreneurs was too small and too insecure to foster the growth of an independent middle class, which might have stood up against the power of the great family combines (the Zaibatsu). Generally speaking, small-scale organization was regarded by many liberals, both foreign and Japanese, as undemocratic, a survival from feudalism. Whether these traits will cling to the small shops under postwar conditions, or whether they can be adapted to democratic aims, is another question considered in the following pages.

II. Prospects for Large-Scale Industry

Quite apart from a possible revival on the part of small enterprise, there is likely to be a relative shift away from the largest units. Both war-time destruction and peace-time occupation policies tend to curtail the large-scale enterprises, at least for some years to come.

The most obvious factor is the effect of the air-raids, which wiped out or badly damaged many large plants which will not be rebuilt in the near future. It is true that the visitor to Japan today sees many big plants (such as the Nakajima aircraft works in the suburbs of Tokyo and the naval repair shops along the waterfront in Sasebo) still standing almost intact on the outskirts of cities where fire has gutted the centers of dense population and small-scale industry. However, a number of great factories were destroyed in the fire raids; others were devastated by precision attacks, which had their successes as well as their conspicuous failures. Furthermore, the damage which was inflicted on large establishments will not be readily repaired, whereas many small, simple workshops can be reconstituted out of diverse parts and odds-and-ends of salvage.

Looking beyond the physical destruction, it is evident that the heavy industries (which are the most highly mechanized and have the largest processing and fabricating plants) are the prime ones

to be removed, eliminated or stringently limited, in order to mini-
mize Japan's war potential and to pay reparations to the Allies.
The manufacture of munitions, including aircraft (which used giant
plants and plant-complexes together with a sizable labor force),
was banned early in the occupation.[9] Munitions plants will be
removed or destroyed; major portions of other large-scale indus-
tries, such as iron and steel, light metals, basic chemicals, heavy
engineering, and shipbuilding, may be made available for repara-
tions removals.[1]

Another important factor is that nations seeking reparations
from Japan prefer to claim the largest and most modern plants —
rather than a multitude of small, diverse plants, many of them
obsolescent, and many designed only for making Japanese-type
goods. The Far Eastern Commission's policy declaration, placing
the Zaibatsu enterprises at the head of the list of plants to be made
available for reparations payments, will certainly have the effect of
earmarking many large plants for removal from Japan.[2] However,
in the case of some civilian-type industries, which are intended to
be the support of postwar Japan, it may prove difficult to remove
all the large plants without destroying the entire industry.·

Other conditions adverse to big business will arise from the
Allies' program for the democratization of Japanese industry. The
program for dissolving the Zaibatsu holdings and for preventing a
future resurgence of these family monopolies may take the form of

9. SCAP Directives of September 6, 1945, September 22, 1945 and
November 18, 1946. (Occupation of Japan, United States Department of
State, Publication 2671, Far Eastern Series 17, page 41).
1. The reparations program which had been proposed by Edwin W.
Pauley, as reported in The Washington Post, November 29, 1946, would have
instituted a drastic curtailment of Japanese heavy industry and would have
hit particularly at the largest plants in certain categories. It now seems
likely that the reparations program will be somewhat less drastic, and will
aim at leaving Japan a reasonably balanced industrial structure. Recent
developments tending in this direction include: (1) the policy decision of the
Far Eastern Commission setting "the peaceful needs of the Japanese people"
at "substantially the standard of living prevailing in Japan during the period
1930–1934" (Press Release, Far Eastern Commission, April 17, 1947);
(2) another decision of the Far Eastern Commission declaring that reparations
shall be in such form as "would not prejudice the defraying of the cost of
occupation and the maintenance of a minimum civilian standard of living"
(The Washington Post, May 21, 1947); (3) increasing attention to the possi-
bility of drawing part of reparations from current production rather than
wholly from capital equipment.
2. "Interim Reparations Removal Program," Occupation of Japan,
p. 156.

placing specific limits upon the size of work force or the scale of operations of any given plant. It is noteworthy that in Germany the American Military Government has decided to break up all enterprises employing more than 10,000 persons.[3] The purge of former industrial leaders, as actual war-criminals or as otherwise undesirable elements, tends to rule out the men who were most familiar with large-scale production, distribution and financing.

The resulting scarcity of big business management will probably be matched by a scarcity of big business funds; for in addition to losses already suffered and the freeze on deposits which is still in effect, recent legislation provides for a capital levy, a property-gains tax, and revision of the income tax rates in the direction of a more progressive structure.[4] These measures will fall heavily upon the great accumulations of capital funds in the hands of a very few persons. For the time being, most of the funds so collected will be impounded in order to combat inflation; later on, part of these funds will probably be spent for public works to cope with unemployment, while another part may eventually become available for loans to business. Lastly, the indemnities which the Japanese government had proposed to pay to business firms as compensation for war damage, and which were an important element in big business plans for postwar reconversion, have been rejected by SCAP headquarters and will not be paid.

Even if it is physically and legally possible to operate some of the large plants, it is bound to be uneconomic and discouraging in many cases. Losses and the fear of losses will certainly arise from the unsettled conditions and the sense of uncertainty in regard to the flow of raw materials and components, the course of prices, the buying power of the domestic market and the access to foreign markets, the release of capital now under freeze, the effect of the proposed capital levy, the designation of specific plants for reparations, the future policies of the Japanese government and the Supreme Commander for the Allied Powers. Large-scale operation is far more vulnerable to most of these fluctuations than a more modest scale of enterprise. Merely the anticipation of great difficulties along these lines will probably keep some of the big plants closed.

3. New York Times, January 2, 1947.
4. See Kurihara, Kenneth K., "Post-War Inflation and Fiscal-Monetary Policy in Japan," American Economic Review, December, 1946.

The number of large-scale establishments will also be reduced by a statistical factor. The use of plant employment figures as a measure of scale of operations means that during the war a very large number of medium-sized plants moved up into the large-scale category as soon as they changed from one-shift to two- or three-shift operation. With the return of peace-time operation, the additional shifts have been dropped in many cases, and the plants correspondingly have dropped back from the large-scale category into the medium-scale or even the small-scale groups.

The textile industry may prove a major exception to the generally unfavorable outlook for large-scale industry.[5] Textiles are a civilian industry. While installed capacity was drastically reduced during the war (by scrapping, transfers overseas, and bomb damage), a substantial fraction of prewar capacity remains in place, in storage, or repairable.[6] Textile products are required both for domestic needs and for export to pay for essential imports. At the same time, the Japanese textile industries are able to utilize raw materials available in other countries, particularly some of the short-staple cotton not suited to the technologies and markets of other producers. The United States has already instituted a program of shipping to Japan substantial quantities of raw cotton for manufacture into fabric and sale abroad. There is a world-wide shortage of textiles today, which is likely to persist for some time to come; in the case of cotton broad-woven fabrics, the quantities available from exporting countries are falling somewhat short of even the minimum import requirements of the various countries, let alone the total volume of demand.[7] For these reasons Japan's textile industry will undoubtedly receive further encouragement to revive. Thus far, it has shown only moderate resilience.

As time goes on, other industries may likewise receive some stimulus. Dean Acheson, as Undersecretary of State, declared that the United States "must push ahead with the reconstruction of those two great workshops of Europe and Asia — Germany and Japan — upon which the ultimate recovery of the two continents

5. It is noteworthy that only the spinning operation is predominantly large-scale. The big spinning companies also own looms in an integrated operation, but these are outnumbered by the thousands of small weaving plants.

6. See The Textile Mission to Japan.

7. Estimates of the Combined Textile Committee, as reported in the Daily News Record, December 2, 1946.

so largely depends."[8] State Department officials are considering ways and means of bringing Japanese productive activity out of its present coma, with major attention going to the importation of much larger supplies of raw materials for fabrication.[9] If some such program eventually goes into effect, the larger-scale establishments may have an opportunity to draw proportionately more assistance from the program than will the small shops, since the former are better suited to a centrally administered plan. However, it is doubtful whether this advantage in itself would enable the large establishments to overcome the numerous barriers and handicaps which have been indicated in the preceding pages.

III. PROSPECTS FOR SMALL-SCALE INDUSTRY

In attempting to fill the vacuum left by the contraction of large-scale industry, the Japanese can resort to medium-scale and small-scale industry. Neither type of enterprise is likely to be opposed by the Occupation authorities, since neither appears to raise any direct conflict with the principal Allied objectives. Medium-sized enterprises (units having from 50 to 1,000 employees) share many of the relative advantages of the small shops, while avoiding some of their disadvantages and weaknesses. On the other hand, medium-scale enterprises will suffer from some of the handicaps of large-scale industry. The prewar and war-time trend of middle-sized units shows a continuous relative decline in their share of industrial employment, but a reversal of the trend is not unlikely under present economic and political conditions. The development of medium-sized enterprises will serve to complement and support small-scale units as well as to compete with them. The former are not likely entirely to supplant the latter, which possess special advantages adapting them to the present situation.

One of the obvious advantages of small-scale industry is its relatively simple requirements for machinery and plant. These requirements often can be met by adaptation of existing facilities. Equipment is still available in towns that were not severely bombed, in the workrooms of farm families, and in rural hiding-places where precious equipment was stored for safekeeping. Even in urban areas swept by fires, it is possible to salvage many pieces of machin-

8. The Washington Post, May 9, 1947.
9. The Washington Post, May 11, 1947, "U. S. Plans 'Hypo' for Lagging Japan."

ery; and the Japanese have shown themselves remarkably adept at putting together diverse parts and piecing out equipment with makeshifts. For these purposes, many of the larger plants which have been damaged or closed can be "cannibalized" and their stocks of spare parts put to use.

In addition to the availability of facilities, the small-scale industries can rely fairly confidently upon a supply of raw materials. Many of the traditional fields for small enterprise, like woodworking, food processing, ceramics, silk textiles, and printing, utilize the products of native agriculture, fishing, forestry and extraction of minerals. Such manufactures will not constitute a heavy drain upon scarce foreign exchange, as would be the case for heavy industries that are dependent upon imports of rubber, petroleum, tin, and a long list of materials almost wholly lacking in Japan. Other fields for small business, such as metal goods, celluloid products, and rubber products, will suffer along with larger enterprises from the severe limitations upon production of the basic metals and the basic chemicals. In these fields, middle-sized plants will often be able to make a relatively stronger bid for the limited available supply than will the tiny shops.

Traditionally the small-scale industries have been devoted to products which are now in great demand for both domestic consumption and export. A great many of the small shops will be able to continue to turn out these goods in the customary fashion. However, some of the shops which served as feeders to large-scale factories will face a difficult problem of adjustment, now that many of the parent factories are closed. In some cases, they will be able to attach themselves to medium-scale plants that will be operating in the same field; in other cases, they may attempt to take over the complete operation in which they formerly performed only a segment — but this attempt is not likely to succeed very well; in still other cases, they will have to change their field of operation. In these adjustments they will be aided by the flexibility characteristic of small-scale industry, which relies in many instances on general-purpose machines, and can shift from one product to another as conditions require.

Small enterprises are well suited in general to cope with the unsettled, uncertain conditions prevailing in Japan today. Drastic changes, which may be brief but recurrent, in the availability of factors of production or in prices or even in government regulations

and taxes, may be sufficient to overwhelm the large firm which has heavy overhead, standardized product and regularized methods, fixed contractual obligations, and substantial inventories. A small enterprise will have better chances of weathering such storms. For example, when a given material is not delivered on time, it is possible to shift to the use of some substitute or to alter the product or change to some other product altogether; or if spare parts are not forthcoming to repair a machine, hand labor can be substituted temporarily or the workers can be diverted to some other stage of the manufacturing process. In such ways, interruptions to production are minimized. When the interruptions do occur, losses are minimized by the relatively small overhead and small inventories, and the personal character of many of the debts. Since labor constitutes so large a share of the costs of production in small shops and capital reserves are so small, such losses as are incurred will usually be absorbed by a reduction in the standard of living of the entrepreneur, his family, and the employees who tend to identify themselves with the enterprise. This system — which is not unknown in other countries — may be deplored as a permanent form of economic organization; but at least the enterprise will remain in existence during a period of fluctuations when many a larger establishment operating under formal profit-and-loss accounting would be driven into bankruptcy.

At the present time, all kinds of production are inhibited by the prevailing situation, and most Japanese are showing a preference for sheer trading — dealing in inventories, black market operations, etc. But when it becomes possible to resume production, small businessmen will face less adverse conditions than will big businessmen, and many an unemployed worker (like many of our own unemployed during depressions) will go into a small productive business on his own.

If we look now at the other side of the picture, we find some serious commercial obstacles to the survival of small-scale units, which are more vulnerable to such obstacles than are medium-sized establishments. To the extent that small shops in the past derived a cost advantage from low wages and poor working conditions, the new protective labor legislation tends to erase this advantage relative to the larger establishments. As regards management, the small units may find it difficult to operate without the administrative guidance and financial support of larger firms.

In the past, many small entrepreneurs were linked with a larger firm or a merchant capitalist in any of several ways already described: sub-contracting, the putting-out system, trade credit, capital loans, lease of equipment, selling arrangements, and so forth. While such guidance and support was in many cases the vehicle of a harsh exploitation, an essential organizing service was performed. Today, many of the parent firms and merchants are inactive or have disappeared. The small firms are thus thrown back upon their own individual devices, or upon devices like the producers' guilds. Producers' coöperatives, with labor union participation in some instances, are being organized. The Japanese Ministry of Commerce and the Economic Stabilization Board have talked of a government program for giving aid to small business.[1] The Japanese government has also proposed to establish a set of control organizations to regulate operations by private owners, but early action affecting small-scale industry is not anticipated.[2]

The problems of management of small enterprises in all countries are well-known. It is unlikely that Japanese small businessmen will be more successful than small businessmen elsewhere in raising funds, obtaining market data, conducting scientific research, rationalizing their methods of buying and selling, instituting effective records and controls — or even in finding out, at a feasible cost, what are the significant questions to ask about their own operations. However, inefficiencies of management, in this sense, are likely to be less serious in Japan during the next few years than in the past, since demand will probably keep ahead of supply, minimizing competition in selling whatever can be produced.

The future of small-scale enterprise as well as the entire industrial organization of Japan may be altered by the trend toward socialization or nationalization of industry, which has made progress in many parts of the world and is gaining adherents in Japan.[3] If nationalization should be permitted in Japan, it would probably be applied gradually, and primarily to those industries whose operations are not small-scale (e.g. coal-mining, transportation, heavy industry, and banking). During this period, the

1. Nippon Keizai Shimbun, July 29, 1946 and September 21, 1946.
2. New York Times, March 9, 1947.
3. See World Report, vol. 2, No. 8, February 25, 1947, pp. 18–19.

socialization program might actually call for aid to small-scale private industry. Should complete socialization of the economy occur, most size-of-plant considerations would become irrelevant.

Balancing all factors, it appears that within the limited volume of industrial activity permitted Japan under the Occupation, small-scale industry can occupy a considerable share. On the one hand, current conditions will stimulate the establishment of many new small units, swelling the numbers persisting from prewar years. On the other hand, the scope and success of their activities will be circumscribed by other current conditions intensifying the traditional disadvantages of small shops. The implications of these prospects are examined in the following section.

IV. IMPLICATIONS

If postwar Japan should find a major rôle for small-scale industry, we may anticipate mixed effects, which will bear on the Japanese economy and on the Occupation program, and will raise questions of social values and social policy.

The most immediate contribution of small-scale industry would be the absorption of large numbers of workers in productive activities. The small shops are labor-intensive, and thus hold out hopes of employment for some of the great numbers of idle workers, offer subsidiary occupations for the underemployed in the farm areas, and provide openings for some of the young people entering the labor force.[4] Furthermore, the small shops will turn out goods when larger enterprises are standing idle, and these goods are urgently needed for Japanese consumption and for export.   .

4. The unemployment and underemployment problem in Japan is very serious. Unemployment, which is not measured very accurately, was estimated at seven million by the end of 1946, or 23 per cent of the labor force. (New York Times, October 21, 1946.) In addition to the figure for that date, about a million more civilians and soldiers were still to be repatriated from overseas, and there is a normal increase in the labor force amounting to about 400,000 persons annually. Early this year Japanese sources estimated total unemployment as high as eight to ten million persons (Asahi Shimbun, February 12, 1947).
   Underemployment on the farms, always a problem due to the extreme density of the agricultural population, is intensified at present by the presence of many persons who fled from the cities during the bombing attacks and who prefer to remain in idleness near food supplies rather than go hungry as well as jobless in the cities. According to the April, 1946, census, 56 per cent of the labor force was engaged in agriculture, in contrast to 42 per cent in 1940. (Summation of Non-Military Activities in Japan, June, 1946.)

From the longer-range point of view, it is possible for small-scale industry to make a contribution toward one of the chief objectives of the Allied Occupation: the development of a strong and stable middle class. This objective is embodied in the "U. S. Initial Post-Surrender Policy for Japan," which states that "policies shall be favored which permit a wide distribution of income and of the ownership of the means of production and trade." [5] This purpose would be served by the expansion of both small and medium-sized enterprises. However, the latter type seems to offer better prospects, in view of higher earnings; more effective organization for operations like foreign trade, and generally greater stability. The small shops can make their contribution to the Allied purpose only if their economic position is stronger and more promising than it was in prewar years.

We have seen that in the past the commercial operations were the greatest weakness of small-scale enterprises. Commercial conditions were such as to raise their costs, to promote cutthroat competition on prices (abroad as well as at home), and to drive wages down. These conditions must be altered if small-scale enterprise is to function successfully. In particular, new financing arrangements are needed, to make capital available to small entrepreneurs at moderate rates, whereas formerly only large enterprises had such access to capital; exorbitant rents levied on small firms must be reduced; new purchasing arrangements are needed to put the individual entrepreneur in a stronger position vis-à-vis the larger supplier of materials and the prime contractor; new marketing organizations may be necessary to reduce the costs and uncertainties of distribution when the product is sold beyond the immediate locality, especially when it is sold abroad (since foreign buyers are reluctant to deal with small firms, and tend to drive hard bargains); adequate channels of information are required, to convey the reports of new processes, new products, new markets; the taxation system needs revision, especially in the direction of permitting the small firm to retain sufficient earnings to encourage rationalization and expansion; subsidies may be necessary to promote these changes, since capital accumulation is feeble, and in general there is need of vigilance to prevent the return of the old desperate exploitative system. If efforts are made to bring about

5. "U. S. Initial Post-Surrender Policy for Japan," August 29, 1945. See Occupation of Japan, p. 79.

these changes, considerable pressure for government aid will probably develop.

In the event that such changes are not made, there is danger that the small enterprises will not be able to rise out of the depressed state in which they were held in prewar years, when small-scale organization was commonly viewed as a semi-feudal relic, paternalistic and anti-democratic. If the general economy of Japan continues at a low level, and if the commercial conditions particularly adverse to small shops are not altered, the independence of these entrepreneurs will be undermined. When numbers of enterprises must look for outside aid, the time would be ripe for a revival of the old financing and controlling system, whose forms and habits will not have passed out of memory in a few years.

The extent and economic health of small-scale industry will also influence the Japanese labor movement, which is encouraged by SCAP as a feature of democratic institutions. Unionization has spread rapidly since the Occupation began, but the solidity of the unions has yet to be tested against the adverse conditions which will rule the Japanese economy. In particular, the large-scale industrial establishments, which are by nature the richest field for union organization, and which in other countries form the backbone of the labor movement, will be severely limited in Japan. Efforts to organize the workers and to maintain them in unions are ordinarily more difficult in the case of small shops than in larger establishments. Furthermore, it is difficult to impose demands for higher wages and better working conditions, in the face of the employer's meager profits, his long hours of work, and his personal contact with his employees. While these difficulties will be strongest in the case of the very smallest units (those having less than ten workers), they apply in some measure to the entire category of small-scale industry.[6]

With respect to the disarmament and demilitarization programs of the Occupation, small industrial units will present a

6. These difficulties have already made themselves felt. While 1,441 collective-bargaining agreements covering industrial workers were drawn up during the period from December, 1945, through December, 1946; only one-third of the agreements were in plants employing less than 100 workers (Summation of Non-Military Activities in Japan, February, 1947, p. 182). This relatively low proportion may be contrasted with the very large percentage of plants falling in the small-scale category (the 1-49 class alone accounted for 96.3 per cent of all industrial establishments, according to the June, 1946, Labor Census).

certain disadvantage. Surveillance of industrial developments will be more difficult to exercise over a multitude of small units than over a correspondingly smaller number of larger establishments. This will be minimized, however, by the nature of small industrial operations, which are usually simple processing; conversion to illegal purposes, such as manufacture of armaments, requires a nucleus of heavy industry and research laboratories, as well as the importation of numerous materials, and these are the critical objects for surveillance.

Another SCAP policy is the promotion of self-sufficiency for Japan by maximizing exports to pay for essential imports. Small shops can be particularly instrumental for this policy, since they are able to utilize a higher proportion of domestic materials and a smaller application of imported machinery than do the larger plants.

Summarizing these implications, it is apparent that the expansion of small-scale industry will make more of a contribution toward the solution of immediate problems than toward the long-run economic and social objectives of the Occupation of Japan. While Occupation policies and economic forces during the next few years will favor both small and medium-sized enterprises over large-scale establishments, the development of an adequate standard of living and a democratic society would seem to call for an expansion of middle-sized industry and a decline of the small shops (particularly those conducted as domestic workshops). However, so long as the Japanese economy generally is in difficult straits, there will be a place for the small shops, because every feasible productive activity will be needed for reconstruction.

EDWIN P. REUBENS.

CORNELL UNIVERSITY

THE CHESTNUT STREET RAID ON WALL STREET, 1839

SUMMARY

I. The business and political situation in 1839, 605. — II. Early sales of postnotes and foreign exchange in the New York and Boston markets, 607, heavy drafts on the New York balances, 609; gathering difficulties in Europe, 609; suspension of the Bank of the United States, 611; the Rothschilds' assistance, 612. — III. Discussion of the episode by Biddle and others, 614; passing of financial primacy to Wall Street, 616; significance of the episode, 617.

In the summer of 1839, a vigorous effort was made by the Bank of the United States, under Pennsylvania charter since 1836, to break the banks of New York. The effort failed. The banks of New York stood, and the banks of Philadelphia were forced to suspend. This settled for good the passing of financial primacy from Chestnut Street to Wall Street.[1]

I

In 1839, business and politics were still in a state of nerves after the panic of 1837. The Van Buren administration, in loyalty to Jacksonian hard money doctrine, was still hoping to order things so that the Government's receipts and disbursements could be effected not only without a central bank but even without use of the private banking system. Chestnut Street, under the inspiration of Nicholas Biddle's later philosophy, was pushing credit into public improvements and national development. Wall Street was contemptuous both of Biddle's program at one extreme and of Van Buren's at the other; under the statesmanlike leadership of Albert Gallatin, she was pursuing a hard money policy that accorded with the best commercial discipline, domestic and international. The Wall Street banks had ended the year-long specie suspension in May, 1838, with little assistance from Washington and despite obstruction from Philadelphia.

1. At the time, the press was full of the story; and Gallatin in his 1841 essay on banking speaks of it specifically. (Gallatin, Writings, III, p. 404.) In 1896, Sumner touched on it in offhand fashion in his banking history. I do not find it mentioned again till another half century had passed, when Sister M. Grace Madeleine in her Monetary Theories of Jacksonian Democracy, published in 1943, referred to it briefly (p. 109). At that time I was working on the subject.

Recognition that her resolute action in 1837-1838 signalized a shifting of the country's financial center was widespread but not universal. Philadelphia still had the now misnamed Bank of the United States, the largest corporation in the country, and she still dominated the South. Nicholas Biddle had made the financing of cotton — the country's principal product and principal export — a main interest of the United States Bank while the latter was a national institution, and the benefits of that interest inured to the successor bank. But in 1839 both the bank and the South in general were still encumbered with slow and over-valued assets, for Biddle had tried to avoid drastic liquidation and had urged others to do the same. He hoped that the economy, with proper monetary management, could be buoyed back into prosperity. He advocated for the country at large what he practiced in the bank. He had first turned to capital investments in anticipation of the United States Bank's liquidation after Jackson's veto, but he, his directors, and his successors remained committed to them after the bank took lease on a new and different life under a Pennsylvania charter. When he retired in the spring of 1839, almost one-third of the bank's earning assets comprised stocks in other banks, in railways, canals, and mines, and bonds issued in the contemporary enthusiasm for public improvements. As the stockholders' investigating committee tardily reported: "So disproportionate an investment of assets in such securities would render any bank essentially unsafe and unsound, and, looking at the fluctuating character of the stock market, make it highly problematical whether an institution thus circumstanced would not be obliged to succumb to the first commercial revulsion that might occur." The bank's loans and discounts were largely secured by similar stocks and bonds. A friendly but judicious critic later commented on the contemporary mania for "mulberry trees, town lots, coal and navigation companies, bank stocks, internal improvements . . . and the thousand delusive speculations" that had misled Mr. Biddle along with the shrewdest; and he expressed amazement at "the folly or ignorance displayed by the bank to the last in making loans to corporations notoriously unworthy of credit."[2]

2. 29th Congress, 1st Session, House Document No. 226, pp. 440, 521. Philadelphia Public Ledger, 4 April 1842. Perhaps it should be emphasized that this later, easy-money, empire-building practice characteristic of Nicholas Biddle's later years was entirely different from the conservative central banking policy he had followed before Andrew Jackson destroyed the Bank

II

In the summer of 1839, three or four months after Biddle retired, sales of postnotes which the Bank of the United States and the Girard Bank of Philadelphia had been carrying on for three years in New York and other money markets were intensified. The postnotes were obligations maturing generally in six months' time and yielding either interest or a discount, or both. They were a form of liability similar to time deposits, but since they could be sold aggressively as investments, they were more readily expansible than deposits. They were issued in denominations as low as $10 and sold at discounts as high as 20 per cent per annum. In Baltimore, it was reported in August: "The whole floating capital of this city has been absorbed in the discounting of post notes of the United States and Girard Banks. . . ." Philadelphia's action "has been of serious injury to us. Now no securities can be discounted at any price." The Boston money market was also invaded; "the bank of the United States put its sucker into their pond in the shape of $800,000 of post notes . . . and from a state of comparative ease, in three days the market was in an agony of pressure."[3] The Boston Courier, 28 August, was very suspicious about the goings on. It reported: "The rate of interest for the past week has been from one to one and a quarter per cent a month for the best paper. This high price for money has been, in a great degree, caused by the large amount of United States Bank post notes, which have been forced upon our market, at one per cent a month discount. The proceeds have been, or are now about to be, taken from our city in specie. . . . Is it not probable, that before these notes are at maturity, the United States and Girard Banks and all others south and west of Pennsylvania may suspend specie payments?"

Meanwhile, the peddling of European exchange was also begun. Money was dear in Europe, and there was no market there for either American products or American securities, whose sale was the usual source of funds abroad. In these circumstances, the balances against which Philadelphia was selling exchange could be

of the United States as a central bank — as I endeavored to show in "Jackson, Biddle, and the Bank of the United States," Journal of Economic History, May, 1947.

3. New York Journal of Commerce, 30 August 1839, 6 September 1839. Boston Courier, 26 August 1839, quoting Baltimore Patriot.

maintained only by shipping specie to Europe or by borrowing there, both of which were expensive. How the bank could be selling exchange so profusely was therefore a mystery. "The United States Bank," reported the New York Journal of Commerce, 30 August 1839, "for sixty to ninety days past has supplied all demands for exchange on England. It has sold several millions of bills at least. What funds these bills have been drawn against for such very large amounts, seeing that American stocks have ceased to sell, we do not understand. . . . And why now the same bank or its most intimate companion is shipping great amounts of gold to England and still drawing bills at a heavy loss compared with the value of the gold it gets out, we are equally at a loss to understand."[4]   In short, the bank was selling drafts on London and Paris, obtaining specie from their sale, shipping the specie to Europe to pay the drafts, and losing money on every draft it sold. "These bills," the New York American later reported, "were pressed on the market with great urgency, and the rates at which they were sold made it matter of easy calculation to ascertain that loss must result from the operation to the sellers, who were to remit the specie they produced in order to meet the bills. As, moreover, the bank was then a borrower in Europe, inquiry was naturally excited as to what these drafts could mean."[5]   According to the Journal of Commerce, the Bank of the United States was "constantly making great sacrifices at somebody's expense for somebody's benefit, neither of which somebodies can exactly be found."   A week later, 6 September 1839, the Journal observed skeptically that "the policy of the bank in drawing bills on England without funds there and shipping specie to meet those bills at maturity is put forth as a voluntary effort on the part of the bank for the public good" . . . , this "good" being the provision of exchange with which merchants could pay their debts in Europe and continue necessary importations.   But however recondite the purpose, the sales were pressed still more hotly, especially of drafts on Hottinguer and Company of Havre and Paris.   The

4. This "intimate companion" was presumably either the Girard in Philadelphia or the United States Bank in New York, a subsidiary organized under New York's Free Banking Act of 1838 by two agents of the United States Bank of Pennsylvania. It was under contract as an agency of the Bank in Philadelphia, which paid all its expenses and a compensation of $12,000 a year.

5. New York American, 20 April 1841.

drafts were signed in blank in Philadelphia and rushed to New York "to be sold without limit."[6]

The sale of notes and drafts yielded the United States Bank substantial balances; and during August it withdrew an estimated $1,250,000 in specie from the New York banks, the largest withdrawals being made in two successive days, the 26th and 27th. On the latter day, agents of the United States Bank, accompanied by notaries, entered various Wall Street banks half an hour before closing time, presented checks of which no previous notice had been given, and demanded instant payment.

"At the Bank of the State of New York, a draft for $80,000 was handed in. The teller said it should be paid as soon as possible. The presenter of the draft replied that it must be paid instantaneously, or the check be given back; that he had no discretion. The cashier being called, and informed of what was going on, said *he* had a discretion in the matter, and that he should *not* give back the check, and that it should be paid by three o'clock. Orders were given accordingly to count out the specie — which was not quite accomplished when the clock struck three. Upon the stroke, appeared the notary, with a duplicate of the draft in the handwriting of Mr. Young, the cashier of the United States Bank here, and prepared to protest it for non-payment. Seeing, however, the money was there, and forthcoming, the protest could not decently be made. . . ."[7]   Similar demands were also met by the other New York banks.

Meanwhile, the lightning was beginning to crackle in Europe, and the United States Bank's agent in London, Samuel Jaudon, was trying to stave off the thunderbolt. On 22 August he wrote for help to Humphreys and Biddle, an American firm in Liverpool. "I must look to you for £50,000, . . ." he said. "If I do not get this I get none . . . everything, therefore, turns upon what you can do, for here I am exhausted. You must therefore work your hardest for me — life or death to the Bank of the United States is the issue."[8]   The following day he wrote for £50,000 more. Next Hottinguer and Company refused to honor the drafts from Philadelphia. Jaudon energetically met the crisis by getting the

6. 29th Congress, 1st Session, House Document No. 226, p. 488.

7. New York American, 20 April 1841; quoted in Niles' Register, 24 April 1841.

8. 29th Congress, 1st Session, House Document No. 226, p. 479. The Biddle of the firm addressed was the son of Nicholas Biddle.

Paris Rothschilds to take them up — or a substantial part of them, but he was still in trouble, and turned to the Bank of England, to which he applied, 26 September 1839 for a loan of £300,000 to the United States Bank for three or four months, to be secured by American stocks. He said in part:[9]

> This application of the Bank has become necessary in consequence of the difficulties which the very unexpected refusal of Messrs. Hottinguer & Co. to accept the drafts of the Bank U. States has occasioned. That refusal was the more unfortunate as my letters of the 31st August from the Cashier of the Bank state that about Fcs 1,800,000 were going to Messrs. Hottinguer & Co. by the Packet of the 1st September and that others would follow in good time.
>
> The discredit however which that refusal has occasioned, notwithstanding that I have made arrangements with Messrs. de Rothschild freres to protect the signature of the Bank U. States so paralyzes my operations that I am compelled to look to the Bank of England as the only power which can sustain me until the arrival of the Steamer Gt. Western on the 5th proximo.
>
> *   *   *   *   *   *   *   *
>
> But there is one point which I may be permitted to suggest, viz., that upon supporting the credit of the Bank U. States depends an operation which, if completed, will enable me to draw upon the Continent to the extent of six hundred thousand pounds and thus bring this large amount to co-operate with the exertions of the Bank of England to turn the exchanges in favor of this country.[1]
>
> *   *   *   *   *   *   *   *
>
> P. S.   In case any loan should be granted by the Bank of England, and it should appear after the Gt. Western arrives that the whole business of the Bank U. States cannot be carried out successfully in the opinion of an eminent House in London, I would make a stipulation that the first remittances from the Bank U. States shall be applied to the repayment of this loan, with the exception of one hundred thousand pounds already advanced to me by a friend who will be prepared, I believe even to postpone this claim in favor of the Bank of England.

The Bank of England refused the application but offered "to advance an amount of Consols of the value of £300,000 for a period not exceeding one month . . . to be secured by the guarantee of Commercial Houses to be approved by the Committee of Treasury." This counter-proposal could give Jaudon little comfort, and it was not accepted. A number of private individuals, however, united to support him and advanced "£800,000 in money,

9. I am indebted to the Governor and Company of the Bank of England for extracting and permitting me to publish the text of this letter.
1. See Clapham, Bank of England, II, p. 168 ff., for an account of the Bank of England's current negotiations with Paris to the same end.

for the period . . . of a year and a half or two years, and upon the deposit of American securities only."[2] These were far better terms for Jaudon in amount, availability of the proceeds, maturity, and guarantee, but the help came too late.

During September, things had been quieter on the western side of the Atlantic, with no repetition of the peremptory performance of August. According to the Journal of Commerce, 31 August, "a letter of conciliation was circulated yesterday . . . from the U. S. Bank full of assurances of good feeling and that no more specie will be drawn and all that." None would be or could be, the Journal was confident, because the postnotes were no longer selling, those sold were maturing, and the balance was the other way. It was, indeed: about six weeks later, 9 October, the Bank of the United States suspended, being quite unable to meet its obligations to the money markets it had been abusing. The other banks of Philadelphia also suspended and the banks of the South and West rather generally followed their example.

The banks of New York and New England did not, though there was stout pressure to do so. The New York Gazette, after rejecting the idea that suspension was necessary or beneficial, found comfort for the New York banks in the assurance that there was now an end "to the postnote suction."[3] The New York correspondent of the Boston Atlas, 12 October, wrote that he had suspected for some time that the Bank of the United States was being "grossly mismanaged"; it had "run wantonly upon embarrassment and dishonor without conceivable motive," had been "borrowing of everybody who would lend it at 15 to 20% in order to lend again . . . at 6", and had been, "in short, making every effort in its power to commit suicide. . . ."

The day after the United States Bank suspended, the steamer from Liverpool arrived in New York with the news that Hottinguer and Company had refused the drafts the bank had drawn upon it, having received neither advice of the drawings nor sufficient funds to cover them. It also brought news of the Rothschilds' help and a report from London to the Courier and New York Enquirer that that help had given the credit of the United States Bank greater strength than ever; the Hottinguers' refusal and the Rothschilds,

2. House of Commons, Report from Select Committee on Banks of Issue, 1840, pp. 149, 231.
3. Boston Courier, 12 October 1839, quoting the New York Gazette.

assistance were "the subject of extraordinary excitement, both in England and in France." Everyone agreed, it was said, that "if Mr. Jaudon, when suddenly called upon, could provide security for 7,000,000 of francs . . . and could substitute the greatest capitalists in the whole world for the comparatively unknown house of Hottinguer — that all this must redound most signally to the credit and solid power of the Bank of the United States."[4] The report mentioned no reason for Hottinguer's action, which was merely presented as extraordinary.

According to Hazard's Register, the Rothschilds' assistance was confirmed in the following letter:[5]

"Paris, 23d Sept., 1839.
"To the President of the Bank of the United States,
          Philadelphia.

"Mr. President: — We have the honour to inform you that we have arranged with Mr. Jaudon to accept for your account the amount of 5,500,000 francs, your drafts on the Messrs. Hottinguer, which remained in suspense. We take it for granted that Mr. Jaudon will have informed you of the arrangements entered into by us with him for this purpose, and consequently consider it unnecessary to recapitulate them here, limiting ourselves to furnishing you, on the other side, a memorandum of such of your drafts as have been left in our hands to-day to be clothed with our acceptance.

"We are happy, Mr. President, to have found an opportunity to give you a proof of our high consideration for the establishment over which you preside, and to have been able, at the same time, to arrest the disastrous effects which this refusal of acceptance on the part of Messrs. Hottinguer was beginning to produce in our place, as well as in Lyons, by many holders of your bills, who, pressed by their necessities to an immediate realisation of their funds, were offering to part with these securities at a loss over the discount.

"We shall correspond with Mr. Jaudon in everything concerning our acceptances on your account, in conformity to his request made to us, so that we shall not be obliged to trouble you with details relative to this operation, except in case of new instructions on your part.

"We present to you, Mr. President, assurances of our most distinguished consideration.

                    (Signed)   De Rothschild, freres,
                               A. De Rothschild."

The Paris Rothschilds came thus warmly and single-handedly to the assistance of the Bank of the United States not two months after they had declined to join a syndicate of Parisian bankers

4. Hazard's United States Commercial and Statistical Register, Vol. I, October, 1839, p. 269.
5. Ibid., Vol. I, November, 1839, p. 335.

formed the latter part of July to assist the Bank of England, whose treasure was fallen so low that convertibility was seriously threatened.[6] Jaudon's success with the Rothschilds, therefore, probably helped him very little in Threadneedle Street. But aside from that and its own straitened position, the Bank of England did not like the ways of the Bank of the United States. The two institutions seem to have had almost no contact till after Biddle was forced by Jackson to give up central banking, had turned promoter, and made the United States Bank a chronic borrower. Then more contacts developed, but with little satisfaction to either bank. In October, 1837, the Bank of England had refused Biddle a loan with the emphatic statement that for her "to furnish capital to any other bank of issue for three years would be *monstrous.*"[7] Yet in 1838 she lent Wall Street £1,000,000 to support resumption and promised Gallatin £1,000,000 more if needed.[8] In this discrimination, the Bank was undoubtedly following the advice of the Barings,[9] which seems to be reflected also in the judgment of The Times, London, that the Bank of the United States was no longer engaged in legitimate banking, but "had constituted itself a great trading and speculating corporation."[1]

With the collapse of 1839, the Pennsylvania Bank of the United States was virtually through forever. She was not suspended merely but "broke," said the New York Journal of Commerce, 12 October, though she did not close definitely till 4 February 1841, after attempting to resume payments and being able to maintain them only about three weeks. Her directors seemed surprised at the failure; despite the generous accumulation of

6. Clapham, Bank of England, II, pp. 168–170. Viner, Studies in the Theory of International Trade, p. 273.

7. Clapham, Bank of England, II, p. 160. Clapham remarks that Governors of the Bank of England did not often use italics. It is interesting that the Bank of England should still consider the Bank of the United States a bank of issue like itself, despite the expiry of the national charter.

8. House of Commons, Report from Select Committee on Banks of Issue, 1840, pp. 117, 154–155; Gallatin, Writings, III, p. 401; Albany (N.Y.) Argus, 16 April 1838. Clapham, Bank of England, II, p. 164.

9. Alexander Baring, Lord Ashburton, was a friend and correspondent of Albert Gallatin, and another firm member, Joshua Bates, was a sort of volunteer adviser to the Bank of England on American affairs. The Barings distrusted Biddle.

1. New York Journal of Commerce, 14 October 1839, quoting The Times, London, 17 September 1839.

reserves they had made in preparing to resume payments, they found that "a feeling of hostility to the institution or, what was equally destructive, a pervading distrust of its credit and means, existed to an extent so great as to render the undertaking hopeless unless the bank was prepared to meet every dollar of her liabilities with a dollar of coin."[2]   Gallatin considered this attempt to resume foredoomed; "it was impossible that it should not have failed," he said.   "The element indispensable for sustaining any bank, *confidence,* was utterly lost.   It seems incredible that it should not have been foreseen that, as soon as the United States Bank paid in specie, every person who held its notes would instantaneously seize the opportunity of converting them into cash."[3]

III

In the spring of 1841, after the third and final suspension of the United States Bank, discussion of the 1839 episode was renewed in consequence of a public statement by Nicholas Biddle that the bank was prosperous when he relinquished management of it, and that its subsequent ruin arose from "efforts to break down the banks of New York."[4]   Biddle's information was that the bank's officers, believing a storm was about to burst, had deemed it best "instead of meeting its full force at once, . . . to make it fall first upon the banks of New York."   This information that the bank had "invited" its own suspension by trying to compel the New York banks to suspend, was something, according to Niles, "which Mr. Biddle is the first to announce."[5]   It is hard to believe that the matter could have been so much a secret, but at any rate it now became common knowledge that in order to get other centers heavily encumbered with demand obligations due to Philadelphia, the United States Bank had plunged into debt itself on obligations of later maturities and had sought by both prolonged and sudden withdrawals to force Wall Street into suspension.   "If the New York banks could have been made to refuse a specie draft," the New York American explained, "the

2. 29th Congress, 1st Session, House Document No. 226, p. 472.
3. Gallatin, Writings, III, p. 405.
4. 29th Congress, 1st Session, House Document No. 226, pp. 480, 487,
5. Niles' Register, 24 April 1841, p. 121.   Philip Hone also indicates that the "precious disclosure" just made by Biddle was news to him.   One would expect Hone to know the facts if many persons did.   Philip Hone, Diary, Allan Nevins, ed., II, pp. 540–541.

alarm consequent upon the notoriety of that fact, would, it was hoped, cause a run upon them that would force a suspension, and then the Philadelphia banks — and especially the Bank of the United States of Pennsylvania, which foresaw its inability to continue specie payments — would have been preceded, and so far justified, in suspension, by those of New York."[6]

Biddle's further statement was that, to effect the bank's purpose, "large means were necessary, and to procure these resort was had to the sale of foreign exchange. . . . The proceeds of these immense sales of exchange created very heavy balances against the New York banks; which, after all, signally failed in producing the contemplated effect. The bills, not being provided for, nor even regularly advised, as had uniformly been the custom of the bank, were dishonored; and although the agent in London did everything which skill and judgment could accomplish, the credit of the bank was gone, and from that day to the present its effects upon the institution have been more and more disastrous."[7] Here, said Biddle, was the "real and secret cause of the disasters of the bank"; the whole trouble was "the neglect, or inadvertence, or omission" to apprise Messrs. Hottinguer and Company of the drafts upon them.[8]

But this is obviously inadequate. It omits to mention the postnotes and it implies that there was nothing behind the bank's action but folly and brigandage: as if its officers, in a state of cloudless well-being, had gratuitously conceived the idea that it would be nice to put Wall Street out of business and to do it by the sale of bills drawn without funds. Yet accounts of the raid followed the emphasis Biddle had put on the sales of exchange, as if those sales had been responsible for most of the demands on the New York banks and as if it were the news of their dishonor that had forced the Bank of the United States to suspend. Since the bank suspended the day *before* the ship from Liverpool arrived in New York, the news of the Hottinguers' action could not have been the immediate cause of the suspension. Moreover, the ship from Liverpool had also brought the news of the Rothschilds' acceptance of the drafts. So it is clear that the actual pressure which closed the bank was that of obligations maturing in the

6. New York American, 20 April 1841; Niles' Register, 24 April 1841, p. 121.
7. Biddle is quoting James Cowperthwaite, former cashier of the bank, who retired about the time he did.
8. 29th Congress, 1st Session, House Document No. 226, p. 488.

domestic market. The jig was up when the raid in August failed and the balances the bank had tried to reverse swung back against it. Niles said explicitly that the bank had been compelled to stop specie payments because of the demands from New York and New England.[9] But it is evident that what captured the interest of less sophisticated contemporaries was Nicholas Biddle's story of the country's largest corporation, which he had once managed, kiting drafts on Paris across the Atlantic Ocean in an effort to break the banks of Wall Street.[1]

The spectacle of "war between the great money marts of Wall Street and Chestnut Street" was fascinating, and there was some tendency to emphasize this rivalry rather than the weakness of the United States Bank as the cause of the raid. It was supported by recollection that Chestnut Street had instigated a meeting at the City Hotel in New York, 23 October 1839, a fortnight after the suspension, for the purpose of inciting merchants and other "honorable men" to embarrass the New York banks with peremptory withdrawals and demands for renewal of their discounts.[2] To this attempt, wrote Gallatin, "the banks, as might well be expected, unanimously refused to yield." If the Bank of the United States of Pennsylvania had been as strong as the Wall Street banks, the episode might be called a matter of rivalry alone, but it is too evident that however bravely she may have thought of regaining her place, the compelling factor in her attempt was the plight she had got into — a plight compounded of false pride and illiquidity. The illiquidity is ironic for an institution which had once been the regulator of the currency and the holder of ultimate reserves.

The passing of financial primacy to Wall Street was the product of many factors, geographic, economic, political, and personal: the Erie Canal, New York harbor, steam-navigation, Jackson's attack on the Bank of the United States, Biddle's later

9. Niles' Register, 24 April 1841, p. 121.
1. "I had a mighty hobby horse
      His name was Nicky-Noddy.
  His head was made of *paper rags* —
    Of cotton bales his body;
  I saddled him, and bridled him,
    And drove to Gotham town.
  There came a "breeze" from *Hottinguer's*
    And blew my hobby down!"
              — Old Nick's Song Book, 1841, p. 5 (Library of Congress)
2. New York American, 20 April 1841; Niles' Register, 24 April 1841, p. 121; Gallatin, Writings, III, p. 404.

credit policy, and the influence of Gallatin. It was also decisive that London definitely preferred Gallatin and New York to Biddle and Philadelphia. For three-quarters of a century after this action, the Bank of England was lender of last resort to the American banking system through the medium of New York banks. The mistake of the Rothschilds in backing the wrong horse so conspicuously at a critical point in the growth of the American economy may explain in part their abstention from American financing thenceforth.

In partial extenuation of the attempt against Wall Street in 1839, it should be made plain that the effort to break banks in a period when notes were still of greater monetary importance than deposits and when convertibility was a universal expectation, had a very different aspect from what it would have now. It was a drastic but well-known form of competition. It meant a stoppage of conversion, if it succeeded, but not the failure of banks necessarily, nor even their closing. There were heavy legal penalties against banks for suspending, but they were not enforceable, especially when suspension was general. In that situation, banks still continued actively in business — even more actively, because the suspension relaxed restraints on their lending and increased the use of their obligations as money. There were plenty of easy money advocates, indeed — followers of Nicholas Biddle and the British anti-bullionists — who thought general suspension a good thing and that the requirement of specie payments curbed the salutary use of bank credit and held back the country's proper development. With a fine mixture of motives and excuses, therefore, the United States Bank doubtless saw advantages not merely to itself if the New York banks could be forced to suspend but to those banks themselves and to the country at large.

I think that to dismiss these ideas as merely excuses for evading responsibility is to miss their substantial historical significance. They involved some reasonable dissatisfaction with primitive hard money doctrine and a pioneer effort to serve the economy with a less inflexible system. To that end, central bank liabilities are now used to supplement inadequate stocks of monetary metal. Nicholas Biddle was already impressed by the way Bank of England notes had served that purpose in the twenty-five years of their inconvertibility, and he was impressed with the wealth of American resources calling for abundant credit. Going beyond the mere avoidance of credit contraction, he sought a

sustained expansion of the economy. To his more conservative contemporaries on both sides of the Atlantic his policy was pernicious — a mere matter of borrowing recklessly and lending recklessly. Albert Gallatin, who in his own words was an "ultra bullionist" and preferred "security to rapid growth," said of the United States Bank that it was "due to the moral feeling of the country not less than to the security of its fiscal concerns that this disgraced and dangerous corporation should not be permitted any longer to exist." The Times, London, while conceding the "patriotic motives and national feelings" that governed the bank's expansionist policy, nevertheless thought its operations "vicious" and that an end to them would be "a benefit to both countries."[3] But it is doubtful if anything of much moment ended with the bank — certainly not that impatience to exploit resources and force the national growth which Gallatin deplored. To that dominant characteristic of the economy's expansion in the nineteenth century, the Bank of the United States was a self-devoted sacrifice, costly and spectacular but not otherwise unique.

In a more restricted aspect the bank's convulsive effort to bring Wall Street to its own level was a clear example of systematic action on bank reserves — a bastard kind of open market operation. Its purpose was to coerce the banks of New York as a group by acquiring claims against them and using those claims to exhaust their specie. The purpose was piratical. Yet the procedure probably looked back to that of former days, when, as regulator of the currency, the United States Bank had restrained the lending of private banks generally. And it looked forward too. A central bank does not make abrupt eleventh-hour demands calculated to be impossible to meet, but otherwise the principle misused in 1839 was that of central banking procedures undertaken nearly a century later for the legitimate purpose of reducing bank reserves and thereby reining in the money market. The episode might also be taken to illustrate the deflationary effect of an import surplus, the exchange offered by the United States Bank being bought presumably in order to pay for imports and the consequence of its purchase being a depletion of bank reserves.

<div align="right">BRAY HAMMOND.</div>

SOMERSET,
CHEVY CHASE,
MARYLAND

3. Gallatin, Writings, II, p. 512, III, p. 406. New York Journal of Commerce, 14 October, 1839, quoting The Times, London, 17 September, 1839.

INTRA-UNION DISPUTES OVER JOB CONTROL

SUMMARY

Scope of the problem, 619. — A. Territorially defined jurisdiction, 619. — B. Jurisdiction defined by classification of work, 621. — C. Traveling employers, 627. — D. Traveling members, 630. — E. Threatened wage standards, 634. — F. Marginal workers, 636. — G. Missionary locals, 638.

The literature of labor economics deals with the dispute over job control almost exclusively as a contest between or among two or more International unions over jurisdictional control of a skill, process, material, or tool. Such disputes, however, are common within a single union. They may grow out of the overlapping of local jurisdictions or out of attempts by locals to expand their jurisdictions at the expense of other locals, or they may result from conflicts of interest groups within the union that are not defined by the jurisdictional lines of locals. Almost invariably they reflect attempts by a group of workers to maximize the area of jobs available to them and to prevent penetration of other workers into this area. Thus, the intra-union jurisdictional dispute is identical in kind to the inter-union dispute. This note attempts to define the circumstances in which job disputes have developed within unions, and illustrates by actual cases particular categories of such disputes.[1]

A. Territorially-Defined Jurisdiction

It is common for International unions to grant to their locals jurisdiction defined by territorial limits. These limits may be

1. Cases were taken mainly from transcripts of union convention proceedings. Most intra-union conflicts are submerged; they occur in dayrooms, union offices, hiring halls, and at union meetings or informally and off-the-record between representatives of contending factions. Most disputes do not reach the conventions at all. Therefore, the cases drawn are clearly partial and the treatment is not comprehensive.

Almost no attempt was made to determine the disposition of disputes, the position of the union leadership relative to the contention, or to indicate which point of view represented challenge to traditional rules of the union or of the industry and which the authority of established practice.

By and large, resolutions which expressed internal disputes are discussed on convention floors only briefly or not at all, and convention decisions in a large majority of cases referred the problem to the union Executive Boards with power to act.

specified generally as a circle marked by a given radius from local headquarters, or as being congruent with county or some other political boundary lines, or they may be spelled out in detail. Occasionally locals are chartered in a given city without further definition of jurisdiction.

Disputes may arise between locals when territorial jurisdictions overlap, when one local seeks to invade the jurisdiction of another, or when two or more locals lay claim to work in an area which is in the territorial jurisdiction of none of them.

Territorial disputes are particularly common in densely populated sections of the country and when a previously inactive area becomes suddenly the scene of building or industrial activity.

1. Local 139 of the International Union of Operating Engineers had jurisdiction over the State of Wisconsin. Local 49, which had jurisdiction over the State of Minnesota, claimed that, when all Minnesota locals were amalgamated into Local 49, they included one local of engineers residing in Duluth, Minnesota but working in Superior, Wisconsin. Both locals claimed jurisdiction over the construction of a grain elevator in Superior.

2. District Council 41 of the Brotherhood of Painters, Decorators and Paper Hangers, in Boston, proposed to expand its territorial jurisdiction to cover the same area as is covered by other craft unions affiliated to the Boston Building and Construction Trades Council. Cambridge Local 577 objected to the expansion of Boston's jurisdiction, its delegate to the union's convention saying, "If this is done, the jobs will be given 100 per cent to Boston men and where will the Cambridge men come off?"

3. In 1935 the Executive Board of the Hod Carriers, Building and Common Laborers considered the question whether Ellis and Bedloes Islands were within the jurisdiction of New York or New Jersey locals. A temporary 50–50 division of work expedient was applied pending promulgation by the Building Trades Department of the A.F.L. of a rule that would apply uniformly to all building crafts.

4. The International Executive Board of the American Federation of Musicians in 1932 considered a dispute between Locals 195, Manitowoc, Wisconsin, and Local 95, Sheboygan,

Wisconsin, over jurisdiction over Strattman's Pavilion, situated midway between the two towns.

5. Streator, Illinois Local 82 of the Hod Carriers complained that Streator is located on the boundary of Livingston County, and that the local union in Pontiac, Illinois was twenty-five miles distant but had jurisdiction over the whole of Livingston County. Local 82 asked that the jurisdictional boundary between it and the Pontiac Local be set at a point midway between the two towns.

6. Local 340 of the Hotel and Restaurant Employees had jurisdiction over bartenders in San Mateo County, California, with the exception of race tracks. Jurisdiction over bartenders at the Tanforan and Meadows Tracks in the county was held by San Francisco Local 41. Local 340 requested transfer. of jurisdiction to itself.

B. JURISDICTION DEFINED BY CLASSIFICATION OF WORK

Some internationals define the jurisdiction of their locals in other respects than territorially. Frequent conflicts develop between locals of specialized classifications of workers in a given locality over the point of division of the work to be controlled by each and between these special-work classification locals in one locality and "mixed" locals in other localities.

1. The Operating Engineers Union chartered stationary engineers locals and hoisting and portable engineers locals. Local 897, Columbus, Georgia, requested and was granted a "mixed" charter, encompassing-both classifications, so that its members might operate hoisting and portable equipment in Muscogee County, Georgia, then within the jurisdiction of Hoisting and Portable Local 312, Birmingham, Alabama. Local 312 opposed change because "hoisting and portable jobs could be more adequately policed by a hoisting and portable local."

2. In St. Johns, N. B., Canada, the International Longshoremen's Association had chartered the following locals: 273 (General Longshore); 810 (Coal Handlers and Bulk Cargo), 1121 (Grain Elevator Employees), 1578 (Scowmen and Lumber Surveyors), and 1039 (Ship Liners and Cleaners). Local 1039 complained to the International convention that

Local 273 refused to accept membership applications from members of other locals during its summertime peak period of work.

3.   Local 345 of the United Brewery, Flour, Cereal and Soft Drink Workers had jurisdiction over bottle beer deliveries in New York.   Local 24 had jurisdiction over draught beer deliveries.   Jurisdictional difficulties arose in the case of mixed loads.   Local 345 proposed that two members should be required on mixed loads, one from each local.

4.   The Hod Carriers' International had chartered Local 310, in Cleveland, as a Building Laborers Local, and Local 330 in the same city as a Building Wreckers Local.   Dispute arose between the two locals over division of work and the International Executive Board ruled that complete demolition is in the jurisdiction of the housewreckers, while all other demolition constituted alteration and came within the jurisdiction of the building laborers.

5.   Hod Carriers' Cleveland Locals (310, Building Laborers; 860, Heavy Construction, and 122, Plumbers' Laborers) disputed jurisdiction over mason tending on brick work done in connection with sewers.   The International Executive Board voted that work done on public sewers and stubs, when the two were laid at the same time, was the work of 860; all work on leads from these stubs or main sewers to buildings belonged to 122; all laying of drain tile, steam lines and conduits, when laid at the time of the foundation by the general contractor, is building laborers' work and belonged to 310; when the work is done by a plumbing contractor, it is the work of 122.

6.   The Brotherhood of Painters, Decorators and Paper Hangers chartered specialized locals of sign painters, and defined the jurisdiction of these locals as "all process work, display work, theatrical and lobby displays, glass tube bending, sign handling, erecting and other necessary operations, also of all sketches, whether made by pencil, ink, paints, or any coloring material and of any painting pictorial work or lettering done on a sketch or sign."   Locals of scenic artists objected to the inclusion of "display work" in the sign painters' jurisdiction.   The scenic artists claimed this work. They complained that scenic artists have been reduced to

straitened circumstances by the destruction of vaudeville by the movies, and they claimed lobby display work and "creations."

7. Bakery and Confectionery Workers' Local 51 has jurisdiction over cake baking in New York City. The Local urged that its jurisdiction be protected in the event of the amalgamation of bread baking locals in New York.

8. Hod Carriers' Local 363 had jurisdiction over Minneapolis civil service workers; Local 563 over private contractors. The City of Minneapolis had bid for a PWA sewerage contract in competition with private bidders and had been awarded the contract. Local 563 claimed this work on the ground that when a political sub-division competed with contractors in bidding for work projects, it is in no different position than any other contractor.

9. The Hod Carriers granted charters to Heavy Construction Locals and Building Construction Locals in the same territorial jurisdiction. Building Construction Locals complained that the Heavy Construction Locals impinged on their work, and asked that they be confined to rock drilling and blasting on roads, viaducts, bridges, etc. and the "heavy concrete construction of same."

10. The Operative Plasterers and Cement Finishers defined the work of cement finishers as "the running of base not exceed six inches in height when same is installed in connection with floors, etc " On a housing project in Houston, Texas, in 1940 cement finishers Local 681 walked off the job because plasterers of Local 79 were laying 6½-inch base. The cement finishers claimed that the original specifications of the project had called for base of 5½ inches with a permissible tolerance of one inch. They claimed that the plasterers had exerted influence to have the specifications read 6½ inches.

11. Plasterers and cement finishers in Buffalo disputed over the laying of base and its finishing to a smooth finish. The work was awarded to the plasterers because it was being done with hawk and trowel, tools whose use is permissible only for plasterers, among union members.

12. Continuing and widespread disputes developed within the Operative Plasterers and Cement Finishers as a result of the introduction of new materials and methods. Main points

of contention were over applying a cement wash to certain types of construction; renovation or restoration of the exterior of a building; and waterproofing of both new and old buildings.

13. Shophand members of the Plasterers and Cement Finishers are ordinarily prohibited from doing scaffold work. Shophand's work consists largely of casting such objects as moldings, cornice work, mantels, brackets, etc. For the past twenty years or so there has been little demand for these products and shophands locals have sought a ruling that it is permissible for shophands to "set up staff," which is work done from the scaffold. The shophands admitted that this was plasterers' work, but asked special consideration on the ground that shophands were old men who knew no other trade and whose trade had been destroyed by changed building specifications. The Southern District Council of the union had permitted shophands to do this work on jobs of less than $500, but an International Vice-President had objected and the decision had been set aside.

14. Local 338 of the Bakery and Confectionery Workers had exclusive jurisdiction over the baking of "beigels" in New York City. The local complained that it had become common practice for members of bread-baking locals to make "beigels." The practice is particularly common among retail bakeries that sell "beigels" in small quantities to grocery stores and delicatessens in the vicinity. Because the practice is so dispersed, Local 338 said it found it impossible to police its jurisdictional rights. It proposed that bread-baking locals should prohibit the baking of "beigels" by their members.

15. The Bakery and Confectionery Workers had twenty locals in New York City, whose jurisdictions were differentiated by products made. Local 1, New York, proposed to the International Convention that amalgamations be fostered to eliminate jurisdictional disputes, since the mechanization of the industry has standardized production processes so that craft differences, by product, have been eliminated.

16. Glaziers Local 1095 of the Brotherhood of Painters, Newark, New Jersey, protested the jurisdiction of Painters' Union mixed locals over glaziers. It proposed that where a special craft glaziers local existed, it should have jurisdiction over all glaziers within a radius of thirty miles.

17. The Longshoremen had occasionally chartered separate locals of longshoremen proper and of pier and dockmen. The Executive Board of the union had then been faced with the problem of settling disputes over the division of jurisdiction between the two.

18. The Boilermakers, Iron Shipbuilders, Welders and Helpers had chartered separate locals of welders. Welders' locals then claimed jurisdiction over all work involving welding. Boilermakers' locals, on the other hand, insisted on doing welding on the ground that American Federation of Labor policy was to regard welding as a tool over which no calling had exclusive jurisdiction, just as no craft has exclusive jurisdiction over the hammer and saw.

19. Local 38, Waitresses, Fresno, California, of the Hotel and Restaurant Employees' Alliance complained to the union convention that Local 566, Bartenders, Fresno, had sponsored a resolution in the California Federation of Labor convention urging state legislation barring women from "mixing, pouring and serving of drinks containing alcoholic liquor from behind a bar."

20. Lead burners (welders) of the Plumbers and Steamfitters sought to have the union constitution changed to require that "all lead construction, i.e., sheet lead, pipe lead and homogenous lead linings that specify lead burned seams," should be done by lead burners exclusively. The lead burners complained that since they did not ask to do plumbing and steamfitting, plumbers and steamfitters should not be permitted to do lead burning.

21. Local 107 of the Pulp, Sulphite and Paper Mill Workers objected to efforts of other locals to organize paper die cutting shops in New York City, over which the Local claimed exclusive jurisdiction.

22. The Plumbers and Steamfitters have a single local (Local 669) with exclusive jurisdiction over installation of all prefabricated sprinklers in the United States. The Local has auxiliary locals in various areas where a sufficiently large number of sprinkler fitters are located. The Newark, New Jersey auxiliary of Local 669 protested the invasion of their work by other classes of locals of the International Union.

23. Local 154, Boston, of the Hod Carriers and Building and

Common Laborers objected to the granting of universal
cards, permitting any work within the jurisdiction of the
International. Local 154 limited its members to plasterers
tending and sought to bar members of other locals from this
work.

24. House Wreckers' Local 95, and Mason Tenders' Local 59
of the Hod Carriers, Building and Common Laborers, both
of New York, disputed over demolition work. The Inter-
national Executive Board assigned complete demolition work
to the Wreckers and alteration work to the Mason Tenders.

25. Jurisdictional disputes developed among New York
Locals 45, 102, 250, 266, and 731 of the Hod Carriers. In 1938
an agreement was signed among these locals assigning juris-
diction. Locals 45 and 102 and any local resulting from an
amalgamation of these two (they later amalgamated to form
Local 147) were assigned "all work done in compressed air;
all lagging, bracing, sheeting, shoring, underpinning"; Local
250 was assigned "all rock drilling and rockdrillers' helpers
known as 'surface work' and open cut and free air shafts
and tunnels"; Local 731 was assigned "all laborers' work done
on the surface and in open cut work." Additional specifica-
tions were made with respect to jurisdiction on work done
partly in compressed, and partly in free air. In February,
1939, the International Executive Board terminated the 1938
agreement and directed the International President to take
over the affairs of Locals 147, 250 and 266. The International
had attempted to sponsor the amalgamation of these locals
and they had consented to amalgamate only on condition
that the amalgamated local should have jurisdiction over
compressed air work throughout the United States and
Canada, all free air shafts and tunnels in the same area, if
safety measures were used, and over an extended area for
tunnel and subway work, drilling, chipping, spading, lasting,
shoring, and timbering, sheeting, lagging, bracing, and under-
pinning work. These conditions were denied by the Inter-
national and the amalgamation did not occur. Thereafter
Local 147, making the general claim to compressed air work
on the ground that members of other locals were inexperienced
in this work and would suffer injury and loss of life and that
other locals could not adequately police this work, engaged

in a series of disputes with other locals and with the International over work in Massachusetts, New Hampshire, New Jersey, Pennsylvania, and in New York State, both within and outside New York City. A dispute with Local 29 (Locals 250 and 266 amalgamated) over work on the Queens Mid-Town Tunnel operation in New York resulted in action against Local 147 by the International, over which there was considerable litigation. Local 17, Newburg, New York, complained that Local 147 interfered with its jurisdiction over heavy engineering construction, and said that Local 147 harassed contractors working in Local 17's territorial jurisdiction by demanding control of and jurisdiction over jobs. The Mason Tenders' District Council of the Hod Carriers in New York complained that Local 147 had picketed its jobs, claiming jurisdiction. Local 60 said that Local 147 had interfered with its jurisdiction on the Delaware Aqueduct job in upper New York State and had engaged in mass picketing on the job. Local 147 complained, on the other hand, that Local 60, exercising complete jurisdiction over the project, refused to honor cards of Local 147 members, although it was accepting non-union men for work on the project; it said that Local 60 was inexperienced in this work and could not police it properly, as a result of which at least twenty-four men had been killed on the job; it urged that Local 60 be deprived of jurisdiction because standards built up by Local 147 for this type of work were being destroyed.

26. Plumbers' and Steamfitters' locals of the United Association of Plumbers and Steamfitters had disputed for many years over division of work. (In Boston they disputed over pipe installation in laundries. Plumbers claimed the work because a laundry simulated a domestic function over which they had jurisdiction; steamfitters said it was commercial work which they were entitled to do.) To eliminate disputes it was proposed that the union revoke previous charters and charter three classes of locals: building trades, metal trades, and combination.

C. TRAVELING EMPLOYERS

Some operations require high-cost equipment and specialized managerial know-how for which there is insufficient demand in a

single locality. To warrant the organization of continuous firms offering services such as these, it becomes economically necessary for equipment to be mobile and for the firm to search for business at points that are geographically widely dispersed. There is constant contention between locals of a craft hired by such traveling employers in whose jurisdiction the work is to be done and locals having jurisdiction in the employer's headquarters' area and covering his regular employees, over the proportion of men that may accompany the employer to the job.

Some Internationals have rules requiring that fifty per cent, or some other proportion, of the men on the job must be taken from the local having jurisdiction where the job is located. Large locals in cities where contractors are headquartered constantly attempt to exert pressure to have this proportion revised downward; smaller locals attempt to have it revised upward.

1. The rules of the Brotherhood of Painters, Decorators and Paperhangers in 1937 permitted a contractor moving into the jurisdiction of another union to take fifty per cent of his own men with him. (The 1937 convention reduced this proportion to twenty-five per cent.) Lansing, Michigan Local 845 complained that some contractors were hiring men who were not members of any local union, having them sign applications, and sending them into other jurisdictions. Local 845 objected to this practice when locals in the area in which work was to be done had unemployed members, and proposed that the application of the fifty per cent rule be made optional with the locals having territorial jurisdiction, so long as they could furnish men for the job.

2. The 1941 convention of the Operative Plasterers' and Cement Finishers' International rejected a proposal that a traveling contractor should be permitted to carry fifty per cent of his men with him into another jurisdiction.

3. The sign painter members of the Painters, Decorators and Paperhangers obtained (1937) a modification of the union rule requiring the hiring of fifty per cent of local men by a contractor doing work in another jurisdiction than his headquarters. The International convention accepted the proposition that the 50–50 rule did not apply to sign painting, if the job could be completed in less than sixteen hours. The sign painters·

said that contractors do not carry house painters with them but bring in outside men only when they secure large jobs, such as hospital or hotel painting. Sign painting contractors, however, get out-of-town jobs that can be completed in a few hours, and the sign painters said that it was wrong to require that this work be split with local members. The Chicago sign painters' local, particularly, complained that it had had difficulty with locals in the vicinity of Chicago which laid claim to half of the work in their jurisdiction.

4. Five locals of the International Union of Operating Engineers had been amalgamated to form a single local with jurisdiction over all of Ohio. Local 114, Cincinnati, had objected to the amalgamation. The General Executive Board rejected the objection, and cited a letter from a contractor expressing approval of the amalgamation because it no longer was necessary for him to change operating crews as he moved from job to job over the State. The contractor claimed that this reduced accidents and equipment breakdowns resulting from inexperienced crews.

5. Licensed Tugmen's Local 374 of the International Longshoremen's Association operated tugs on the Great Lakes. It sought to have its men take over barges being towed through the New York State Barge Canal by members of United Marine Division Local 333 at the Canal Lake Terminal.

6. New York Local 8 of the United Slate, Tile and Composition Roofers, Damp and Waterproof Workers' Association proposed that fines be imposed on locals which did not inform the local with jurisdiction in a firm's headquarters area when firm entered the territory of another local, "so that the interests of the men in the firm's local should be protected." Local 8's contracts provided for fifty per cent use of Local 8 men on out-of-town jobs. The resolution was intended to enforce this contract clause. Out-of-town locals who would lose fifty per cent of the work done in their jurisdiction by New York contractors by enforcement of the clause objected, and said that Local 8 should police its own contract.

7. Local 12 of the International Union of Operating Engineers claimed jurisdiction over dredgemen in Southern California. Local 3, located in San Francisco, claimed that

ninety per cent of Southern California dredging was done by San Francisco contractors. Local 12, on the other hand, said that Local 3's headquarters were too far away from Southern California for the work to be adequately policed.

8. United Automobile Workers-CIO Graham-Paige Local 142 and Willow Run Local disputed for jurisdiction over the Willow Run plant, in Michigan. Local 50 had had jurisdiction over airframe operations of Ford Motor Company in the plant. Local 142 had had jurisdiction over the Graham-Paige operations in its Detroit plant. The Ford operations had been abandoned and the plant had been taken over by the Kaiser-Frazer organization into which Graham-Paige had merged. Local 142 claimed exclusive jurisdiction, so that Kaiser-Frazer hiring would be done from its seniority lists. Local 50 urged the consolidation of the seniority lists of the two locals. The dispute was complicated by the fact that the two locals were associated with opposite political factions in the union.

9. Manhattan longshoremen members of the International Longshoremen's Association refused early in 1947 to unload a ship for which the American Export Lines was agent, because the shipping line had diverted many of its heavy cargo vessels to Jersey City and Brooklyn for unloading.

D. Traveling Members

Internationals covering crafts of which a significant number of members travel are faced constantly with the problem of disputes between traveling members and "stay-at-homes" over job rights. Non-traveling members, especially where locals have been successful in controlling the supply of craftsmen in their jurisdiction, object to the "invasion" of traveling members. This is especially true in periods of unemployment. In some cases, locals are permitted to reject transfer cards of members of other locals, when a specified percentage of the local's own members are unemployed. If a local's membership includes a specialized sub-craft, this group sometimes seeks to have the percentage-of-unemployed rule applied by sub-craft specialty, rather than to the local's total membership. Sometimes disabilities are imposed on the traveling member or on employers of traveling members in efforts to expand the employment opportunities of resident

members; more or less ingenious subterfuges and circumventions are then developed by traveling members to evade compliance with the union rules.

1. The law of the International Photo-Engravers' Union permits a local to refuse to accept traveling cards of members of other locals when ten per cent of the members of the local into which the move is made are unemployed. The roto-gravure workers sought to have the ten per cent rule apply by sections of the craft; thus, if ten per cent of rotogravure worker members of a local were unemployed, the local could refuse traveling cards from "foreign" gravure workers, even though less than ten per cent of the total membership of the local were unemployed.

2. The 1940 convention of the Brotherhood of Teamsters, Chauffeurs, Stablemen and Helpers made it compulsory, by constitutional amendment, on locals to accept transfer cards of members of any other local. Delegates objected to the taking of new members when old members were still unemployed.

3. Tacoma, Washington Local 252 of the Hod Carriers, Building and Common Laborers complained that the International Constitution permitted the issuance of traveling cards to members immediately upon being sworn into the union. This permitted locals to admit applicants, benefit from receipt of initiation fees, and for traveling cards to be immediately issued, so that the local admitting new members did not suffer, but locals to which traveling members moved, did. The Tacoma local proposed that the Constitution be revised to permit the issuance of traveling cards only to members in good standing for at least one year.

4. Theatrical Wardrobe Attendants' Union No. 16770 of the International Alliance of Theatrical Stage Employees petitioned that traveling members of the Alliance be prohibited from taking their wives on tour to act as wardrobe attendants.

5. The problem of the traveling member has been a particularly serious one for the American Federation of Musicians. Already faced with unemployment resulting from the expansion of mechanical transcription, radio broadcasting, and

sound motion pictures, the depression after 1929 played
havoc with employment opportunities of members. To pro-
tect employment of local members, the union, during the
depression, required that traveling bands charge thirty per
cent more than the local price when carrying out an engage-
ment in another local's jurisdiction. Circumventions of one
kind or another developed to evade the requirement, with
traveling bands sometimes entering into collusive arrange-
ments with their employers. Some traveling members took
engagements in neighboring towns and returned home each
night, claiming that this constituted a succession of one-day
stands, which were exempt from the thirty per cent rule.
Sometimes the thirty per cent over local price was charged and
"kicked back" to the employer. Sometimes double contracts
were made, a dummy, incorporating the thirty per cent extra
charge, and an actual contract for a lesser price. Sometimes
the thirty per cent overcharge was borne by traveling members
themselves, rather than by the employer. The International
officers urged the revocation of the thirty per cent law because
it protected no one's employment and invited contempt for
the union's laws. This was done by the 1934 convention
which required, in its place, a ten per cent charge by traveling
orchestras over the local price. The ten per cent law, too,
came to be evaded when traveling orchestras would take
membership in a small local close to the place of their engage-
ment, claim this as their home local, and claim exemption
of their engagements on the ground that they constituted
"local out-of-town" engagements, which were exempted from
the law. To combat this practice the International Presi-
dent's office struck out the exemption of local out-of-town
engagements, but did continue the exemption of private
functions of lodges, societies, etc. Thereupon, traveling
orchestras began to claim all engagements to be private.
Therefore, a new law was promulgated by the International
President striking out all exemptions.

6.   The constitution of the Plumbers and Steamfitters' Union
provides that clearance cards be issued only to first-class
journeymen. Traveling members depositing clearance cards
with another local may not be examined for skill qualifications
or be required to pay any fee. During the war, many helpers

and apprentices traveled to war-plant construction jobs. It was proposed to the 1941 convention of the union that these members also be issued clearance cards.

7. Local 655, Miami, Florida, of the Musicians' Union asked special treatment with respect to traveling members coming into its jurisdiction. It claimed that traveling members were one and one-half times as numerous as the local's permanent membership and that, therefore, the resident members suffered greatly from the competition of outsider members. The local claimed that some out-of-town members accepted non-union engagements in Miami at below the local scale. It also claimed that some members accepted engagements with irresponsible promoters, knowing that the enterprise would fail, so that they might then search for other jobs within the jurisdiction. The local asked for a law which would require traveling orchestras to hire resident members in the number of fifty per cent of its total, and that the local be permitted to reject transfer cards in numbers greater than fifty per cent of the local's permanent membership.

8. In the early years of the Musicians' Union, theatre musicians' wage rates were very low, and the International encouraged locals to enforce higher rates. Many locals were successful, but theatre managers, forced to raise their rates, retaliated by hiring orchestras from another local. Therefore the Federation enacted a law providing that members could not solicit or accept a theatre season engagement in the jurisdiction of another local, except with the consent of the local having jurisdiction. This law was later extended to radio broadcasting for the same reason. However, when big-name broadcasting bands developed as an advertising technique, it became the practice for a local to consent to the playing of a traveling orchestra within its jurisdiction, on the condition that an equal number of standby members of the local should be hired. However, one local which, according to the International President's report to the 1933 convention of the union, had "an exceedingly hateful disposition toward every musician who ever transferred into the local," refused to permit traveling musicians to broadcast from its jurisdiction, even though the standby-hiring conditions were fulfilled. Therefore the law was changed to require that a

local permit radio broadcasting from its jurisdiction by outside orchestras, provided that the standbys were hired from the local.

9. The 1934 convention of the Musicians' Union adopted a law which prohibited a member accepting an engagement with a symphony orchestra in another jurisdiction at $40.00 or more per week from accepting single or other engagements in that jurisdiction of a non-symphonic nature without the consent of the local having jurisdiction.

E. Threatened Wage Standards

Disputes sometimes arise between locals of an International when firms "run away" from the jurisdiction of a higher- to a lower-wage standard local. In these cases, the higher-standard local from which the firm fled may claim jurisdiction over the runaway itself, or it may seek to have its jurisdiction expanded to include the area in which the runaway firm locates. In other cases the product of a high-standard local may be faced with the competition of products produced in lower-standard areas. Here the high-standard local may lay claim to jurisdiction over the area where the competitive product is produced, and claim the expansion of jurisdiction to be essential to the protection of its members.

1. Window Cleaners' Protective Union Local 2, New York, of the Building Service Employees' Union was chartered in 1916, and until 1936 represented all union window cleaners in New York. In 1936 window cleaners who were self-employed were chartered in Local 95 and subsequently additional charters for self-employed window cleaners were issued (Locals 156 and 115). Local 2 complained that its wage and hour standards were being destroyed by members of the locals of the self-employed, who worked no set hours and for no set wage. Contractors hiring Local 2 men were said to be losing business to the self-employed. Local 2 said that when the self-employed locals were chartered their members were few in number and worked in small jobs. Since then, however, they had begun to invade the craft of commercial window-cleaning, and Local 2 asked the cancellation of the charters.

2. Locals 11 and 12 (Danbury, Connecticut) of the United Hatters, Cap and Millinery Workers requested expansion of

their jurisdiction to cover Bethel, and that the charters of Locals 61 and 62 (Bethel) be revoked. The Danbury locals urged this course of action because low standards in force in the Bethel Plant were threatening the enforcement of Danbury standards in their jurisdiction. Locals 61 and 62, on the other hand, said that the Bethel shop had threatened to shut down, if it were compelled to adopt Danbury standards.

3. The Norwalk local of the Hatters objected to the employment of women in Fall River plants in the sizing and starting departments, which involves working at steaming vats. The local proposed that legislation be introduced into the Massachusetts legislature prohibiting the employment of women in these departments, on the ground that it was deleterious to health. The probable real objection of the local was based on the competitive threat to the wage standards of its members which employment of women involved.

4. Local 754 of the Hod Carriers, Building and Common Laborers had jurisdiction over Rockland County, New York from its charter date in 1926 to 1938. In that year Local 17 was chartered and given jurisdiction over heavy construction work in this territory. Local 754 complained that Local 17 had a wage scale of 65 cents per hour, in contrast to its own scale of $1.00 per hour. Local 754 said that its standards were being destroyed by Local 17 and requested that it have complete jurisdiction over Rockland County.

5. The United Hatters, Cap and Millinery Workers had two millinery locals in Chicago, Local 51 composed of male members and Local 52 composed of female members. Local 51 proposed that the two be amalgamated, because the split "conspired to defeat the unity and solidarity of the workers" and because manufacturers lowered wage standards by using the split to foster wage competition.

6. Local 1 of the Bakery and Confectionery Workers controlled the entire frankfurter roll industry in Greater New York. In 1939 the Sabrett Food Products Company moved from New York to Jersey City and there signed a contract with Local 256. The contract did not provide for renegotiation in the event of mechanization. Sabrett did mechanize and increased its output at the same wage rates, shipping its product into New York. Local 1 charged that the employ-

ment of one hundred frankfurter roll makers in New York was jeopardized by this competition, and asked that it be granted jurisdiction over this Company.

7.   The Bricklayers', Masons' and Plasterers' International was faced with the problem of whether to admit into membership in the union persons working on real estate construction and allow them to accept a secondary wage rate which would be less than the rate on heavy construction.   International officers sponsored this proposal, because the housing field could not be organized in any other way.   The proposition was opposed by the New York local, which feared that the differential would result in deterioration of wage standards on heavy construction jobs.   The International officers were supported by locals in cities where there was little heavy construction and in which union members had to find jobs in housing.

8.   Local 311 of the United Brewery Workers had jurisdiction over seltzer shops.   Since the repeal of prohibition, many soft drink shops in New York had begun to produce seltzer at a daily wage rate (1939) of $7.00, compared with Local 311's rate of $9.00.   The local proposed that only its members be permitted to produce seltzer, to protect its wage scale.

9.   The Hotel and Restaurant Employees had three waiters' locals in New York City, Local 1 having jurisdiction from the Battery to 42d Street, Local 16, from 42d Street to 112th Street, west of Fifth Avenue; and Local 219, from 58th Street to 138th Street, east of Fifth Avenue. Members were restricted to work within the jurisdiction of their own local.   Local 16 complained that there were disparities with respect to wage standards and conditions of work among the locals and that night clubs were able to "run away" across the street to lower scales and standards.   Locals 16 and 219 were on record in favor of amalgamation, but Local 1 was opposed.

F. "MARGINAL" WORKERS

It is a common practice for whole categories of union members, often not defined by local lines, to lay special claim to job rights. The special claimants may be a particular race, workers with longer terms of service in the craft or industry or with their employer, permanent year-round workers as distinguished from

seasonal workers, male workers, journeymen, or workers qualified in the all-round skills of the craft. Workers outside the category are then in one way or another limited in their access to jobs.[2]

1.   Local 844, Seattle, of the Hotel and Restaurant Employees had jurisdiction only over Orientals employed in Oriental establishments. It also sought to organize Orientals in Occidental establishments. The local had done this because Cooks' and Assistants' Union 333, Seattle, had refused to admit Orientals to membership. Local 844 sought permission to extend its jurisdiction to cover Orientals in Occidental places of business and to furnish Orientals to Occidental establishments from its hiring lists, provided they worked under the standard union agreements. This was rejected by the convention.

2.   In 1933 the International Executive Board of the American Federation of Musicians considered the question of which of two locals in a city, one white and the other colored, would police traveling colored bands coming into the jurisdiction. The Board decided that either should have power to investigate and summon members, but that trial for alleged offenses should be at the hands of the Board.

3.   The International Longshoremen's Association chartered separate white and colored locals in Southern ports. Disputes constantly developed over division of work between locals.

4.   Southern Negro locals of the Brotherhood of Carpenters and Joiners complained that their referrals to jobs were rejected because members of white locals in the same locality refused to work with Negroes.`

5.   St. Louis members of the United Hatters, Cap and Millinery Workers objected to granting peak-season extra workers permanent status in the union and a share in off-season work. The St. Louis local requested an amendment to the constitution to provide for the issuance of temporary cards. The convention committee rejected the proposal on the ground that it must be a general proposition of the

---

2. The cases of Negroes or other minority races being excluded from union membership, of superior job rights vested in classes of workers with high seniority ratings, and of limitations on the number and job rights of apprentices and helpers are excluded here because they are widely practiced, well-known, and have been adequately treated in the literature.

union that all workers coming into the union have equal protection, and because workers given temporary cards are unorganizable and tend to take year-round jobs in non-union shops. The committee recommended that locals that find the general rule burdensome seek special relief through consultation with International officers.

6. Subordinate unions of the Bricklayers', Masons' and Plasterers' International resisted the policy of the National Office of the union to admit cleaners, pointers, and caulkers to membership. The National Officers took the position that men not qualified as journeymen bricklayers should be admitted as specialists in a sub-segment of the craft and held to this type of work. They claimed that, if not admitted, these men would work under non-union conditions. Bricklayers' locals, however, complained that admitting these men would deprive qualified bricklayers of caulking and pointing jobs. They also objected that men admitted to the union and granted caulking cards had sometimes carried trowels and done patching work belonging to bricklayers.

7. The 1946 convention of the International Brotherhood of Bookbinders voted that women bookbinders should be paid equally with men for equal work, but stipulated that women should not be hired if men were available.

8. Local 125 of the International Fur and Leather Workers' Union consisted of floor boys. The local sought to affiliate to the Joint Council of Furriers in New York, so that its members might have larger access to mechanic furrier jobs. Mechanic furriers, who largely followed the practice of apprenticing their sons in the craft, resisted Local 125's demands.

G. Missionary Locals

In some instances locals take an aggressive interest in the organization of workers of their craft. Sometimes before and sometimes after the local has been able to complete the organization of workers covered by its nominal jurisdiction, it will undertake to organize workers who are in the jurisdiction of another local which has been disinterested or, for other reasons, inactive towards organizing these workers. Disputes develop when the latter then claims jurisdiction.

1. Local 17 of the Bakery and Confectionery Workers (Astoria, Long Island, New York) said that it had sought to organize the industry in its area, but that locals having contiguous territorial jurisdiction had objected to infringement of their jurisdictions, and that this had slowed down the organizing drive. Local 17 proposed that it should be permitted to organize bakeries "adjoining [its] territory," if the local union having jurisdiction was unable to organize them or did not do so.

2. Local 70 of the Upholsterers, Furniture, Carpet, Linoleum and Awning Workers had jurisdiction over floor covering in New York City. It largely ignored workers going out of small shops for work in apartment houses, but these workers organized and affiliated to Local 70 as Local 70-B. Local 70-B requested the assignment to it of jurisdiction over floor covering in Kings, Queens, and the Bronx Counties, and this was granted over the objection of Local 70.

The interests of local unions are as parochial as those of the Internationals with which they are affiliated. Just as for the International, it is a common principle of action for locals to create job monopoly situations in which employers are compelled to hire from among local members and limited numbers of persons are admitted to the local's membership rolls. Each local seeks to expand the job area that it controls. The cases cited in this paper indicate the extent of the problem created by provincial job interests and the pressure to maximize job control within the International Union.

<div align="right">Simon Rottenberg.</div>

Harvard University

STUART WOOD
AND THE MARGINAL PRODUCTIVITY THEORY

Stuart Wood has two claims to economic fame: he appears to have been the first person to receive a Ph.D. in economics in America and he was one of the independent discoverers of the marginal productivity theory. The latter claim, at least, is substantial: he shares this discovery with economists of the stature of Marshall, Wicksell, Barone, and J. B. Clark. I shall first present a few biographical details and then summarize Wood's economic writings.

I. Biographical Details

Stuart Wood was born in Philadelphia on May 30, 1853, and died there on March 2, 1914.[1] His father, Richard Davis Wood, was a prominent banker, manufacturer, and railroad builder, and he eventually followed into the former two industries.

Wood was apparently privately tutored before entering Haverford College in 1866 at the age of thirteen,[2] where he had been preceded by his five older brothers. His academic record was very good: his annual average grades were 98.0, 98.3, 98.9, and 99.6 — although one should add that grades over 100 were not unknown. The following three years were spent chiefly in travel, including a year of study of philosophy at an unknown German university.

In 1873 Wood entered the Graduate School of Harvard University. He was in residence for two years, and wrote (in longhand) a 133-page thesis entitled "A Review of the 'Principles of Social Science' by Henry C. Carey." In 1875, a month after becoming twenty-two, he received a Ph.D., the first in economics at Harvard. His work in economics, under Professor Dunbar, had included readings of Blanqui's "History of Political Economy," Smith's "Wealth of Nations," and Carey's "Social Science." It was eight years later before a second Ph.D. in economics was conferred by Harvard — on F. W. Taussig.

1. The sources of information are various: obituaries; university records; correspondence with associates; etc. I am particularly indebted to Mr. C. B. Suttle of Philadelphia, who was Wood's secretary during the latter years of his life.

2. His record at Haverford College gives only the Westtown Friends School (at Westtown, Pa.) as earlier training; the records of this latter school indicate that he attended only for five months in 1863.

Even before completing his academic work, Wood had entered his father's manufacturing business. In 1874 he formed a partnership with four of his brothers, to continue as iron founders and machinists (with mills at Florence and Millville, N. J.) and as commission merchants and dealers and finishers in cotton and other goods. He specialized in the financial end of the business. He also became privately interested in several public utilities (Millville Gas and Water Co., Tampa Water Works Co., Macon Gas and Water Co., etc.), banks (Vineland National Bank, N. J., and the Market Street National Bank of Philadelphia), and acquired large acreages of undeveloped coal and timber lands in West Virginia and property suburban to Philadelphia.

He must have had more than average energy, for in addition to these numerous business activities, he dabbled in politics,[3] was manager of the Pennsylvania Institute for Deaf and Blind, treasurer of several learned societies, and active in the American Economic Association. He joined this association in 1886 shortly after it was founded and was a member until his death. He was long a member of the Council, a body much like, and in the course of time not much smaller than, the Legion of Honor.[4] He attended numerous meetings of the Association (at one of which he played host to the entire group), helped finance early journalistic ventures of the Association, and was a vice-president in 1901.

His active writing in economics fell in a short span of time (1888 to 1891). It is reported that only early and extensive business entanglements prevented him from entering academic life, for which, as we shall see, he had talents not common in faculties of economics at that time.

II. Economic Writings

Stuart Wood's doctoral dissertation, it must be reported, was an undistinguished performance. It was a selective summary and a non-selective criticism of Henry C. Carey's "Principles of Social Science." Bits of the analysis were fresh, but on the whole Wood was content to repeat J. S. Mill and his predecessors. The style

3. In 1884 he organized the Independent Republican movement in Pennsylvania to support Cleveland.
4. When he was first appointed in 1886, the Council had forty members; by 1903 this number had grown to 166.

was flowery, the tone superior,[5] and the standard of craftsmanship low.[6] Yet few who, after a decent interval, have reread their own theses will fail to have some sympathy for these shortcomings. And for Wood there was the extenuation of youth — the thesis was written when he was twenty-one.

Our real interest, however, is in his later essays. They fall into the two fields of the marginal productivity theory and the history of doctrine.

The Marginal Productivity Theory. The marginal productivity theory was presented in two essays which appeared in 1888 and 1889.[7] It is assumed in these essays that the supply of labor is fixed, and the investigation then turns on the manner in which the price of labor acts on the quantity of labor demanded so as to lead to an equilibrium with full employment.[8] The fundamental determinant of the demand for labor is immediately stated to be the substitution relationship between capital and labor. Wood is concerned chiefly with two problems: (1) the manner in which one productive factor is substituted for the other; and (2) the determination of the rates of wages and interest.

Substitution of one factor for the other takes place in two general ways. The first method of substitution is inter-industrial:

> All articles are produced by the coöperation of labor and of capital, but it may be that no two articles are produced by exactly the same proportions of them. As consumption happens to be of those articles into whose production labor enters more largely, or of those other articles into whose production the

5. The following passage is an exaggerated example:
"Before examining Mr. Carey's own theory, I now turn to what he says about Mr. Malthus. He attempts in an early chapter to refute him offhand in a very few words,— words which demonstrate either Mr. Carey's own complete and stupid misapprehension of Mr. Malthus' meaning, or else his wilful misinterpretation of it,— which refutes nothing but that silly claim for fairness which Mr. Carey has the impudence to put forward in his preface (pp. 66–67)."

6. Compare the light-hearted accusation of plagiarism:
"Many of Mr. Carey's views, in fact, are closely related to those of List, and although I have not been at the pains to examine with accuracy the question of priority, all the circumstances seem to point very strongly to liberal plagiarism on the part of the former (pp. 51–52)."

7. "A New View of the Theory of Wages," this JOURNAL, III (1888–1889), 60–86, 462–480; and "The Theory of Wages," Publications of the American Economic Association, IV (1889), 5–35. The latter paper was presented at the same meeting at which J. B. Clark first presented his version of the theory, in "Possibility of a Scientific Law of Wages," ibid., 39–69.

8. "New View," pp. 61–63; "Theory of Wages," pp. 6–9.

use of capital enters more largely, just in such proportion does it occasion a large or a small demand for labor.⁹

The technical relationships between capital and labor and the conditions of demand together determine the equilibrium position:

> If the price of labor be excessively high, the price of those commodities into which labor most largely enters will also be high, and the demand for them will slacken. Employers will withdraw capital from their production in order to embark it in industries using a greater proportion of auxiliary capital. The demand for labor and its price will fall, and the use of auxiliary capital will grow at its expense until equilibrium is restored.¹

This method of inter-industrial substitution had already been formulated with greater precision by Walras and Wieser, and later Cassel made it the fundamental part of his theory of production.

The new and important element in Wood's theory, however, is his second method of substitution. It may be termed intra-firm:

> Of the vast number of processes which are carried on by machinery, the greater number perhaps could be executed as well by hand labor. Machinery is preferred simply because the cost of production is lessened by its use. Also, in many processes now executed by hand, machinery would be equally efficient; but it is not used, because it would involve a greater cost of production. In these cases, opportunities for labor or for auxiliary capital to supplant the one the other only occur when either the price of labor or the cost of using machinery changes. In some cases, a great change only will afford such opportunities: in other cases, a small change will suffice. But there are other cases in which the cost is equal, or very nearly equal, whether labor is employed or auxiliary capital. In such cases, the slightest change in the price of labor immediately affects the demand for it.²

Wood comes close to the modern concept of a marginal rate of substitution; he speaks, in fact, of the "rate of interchange" between factors.³ There is a strong temptation to attribute clairvoyance to him when he states that "Labor and capital are interchangeable and are used indifferently where . . . each may be said to be at its final relative utility,"⁴ for the marginal rate of substitution of capital for labor (in the Hicks-Allen terminology) is the ratio of the marginal productivity of labor to that of capital. Again, diminishing returns (to capital and labor) are stated primarily in terms of a scale of the quantities of capital necessary to

9. "Theory of Wages," p. 11.
1. "New View," p. 84.
2. "New View," pp. 67–68.
3. "New View," pp. 73, 463, 476; "Theory of Wages," pp. 19, 30, 34.
4. "Theory of Wages," p. 29 (italicised by Wood).

supplant one laborer in various occupations.[5] With such an approach, it is not surprising to find also an excellent statement of what is commonly called Marshall's "law of substitution":

> In productive industry, every form of investment is open to the employer. . . .
> In all cases, he is governed by one and the same motive — namely, the desire
> for gain; and, according to his expectations as to the comparative profit to be
> reaped, he turns capital into one field or into another. He does not buy labor
> until he has compared the profit to be got by so doing with that obtainable by a
> different outlay. . . . If labor commands a high price, he economizes in its use
> and employs more auxiliary capital.[6]

Let us turn now to Wood's second problem, the determination of the rates of interest and wages. Given the supplies of the productive factors, the tastes of consumers, and the state of technology — the list is his — the wage rate and the interest rate are determined by three conditions: (1) full employment of capital and labor;[7] (2) equalization of marginal utilities of various commodities per dollar's worth of expenditure;[8] and (3) the marginal productivity theorem: "the relative final utilities [marginal productivities] of labor and of capital fix their relative prices, and the wages of labor and the interest on capital are the same in all their employments as they are here." [9] All of the essentials of the theory of general equilibrium are present in this analysis.

An interesting application of the theory is made to international trade.[1] Wood is one of the few nineteenth century economists (Sismondi and Longfield must also be named) to explain the existence and nature of international (and interregional) trade directly

5. "New View," pp. 68, 464; "Theory of Wages," especially the graph on page 31 and the accompanying discussion.
6. "New View," pp 63-64.
7. "New View," pp. 63, 480; "Theory of Wages," pp. 13-14.
8. "New View," p. 84: "Equilibrium can only exist in the price of labor and in the price of using capital when it is impossible to augment the sum of gratifications by supplanting industries employing a large proportion of one by industries employing a large proportion of the other. Among the industries between which desire is evenly balanced, there are some in which labor predominates, and some in which capital predominates. The price of labor which will produce any gratification in the former industries equals the cost of using capital to produce in the latter other gratifications equally esteemed by human desire." See also "Theory of Wages," p. 9.
9. "Theory of Wages," p. 14; also "New View," pp. 63, 66.
1. Considerable attention is also paid to the absolute and relative shares of capital and labor, but the discussion does not get beyond a listing of determinants — a position at which it still tends to come to rest. See "New View," pp. 72-73, 463-478; "Theory of Wages," pp. 17-29.   •

by differences between countries in the relative supplies and prices of capital and labor:

Each country will devote itself to the growth of that crop for which it is best suited, and will supply its wants of other things by exchange. If the people of the country which grows oranges [which require much capital] wish for flax [which requires much labor], they will obtain it by giving oranges in exchange to the people of the flax-growing country. . . . Thus the needs of each country will be supplied in the cheapest manner. The demand for labor, however, will be far greater in the flax-growing country relatively to the demand for capital, and should the supplies of labor and of capital be the same in both countries, wages will be higher where flax is grown and interest where oranges are grown.[2] If in any two countries the relative prices paid for labor and for the use of capital differ, then such industries as require most labor tend towards the country where labor is relatively cheapest, while such as require most auxiliary capital gravitate to the country where labor is relatively dearest and where interest is lowest.[3]

We need not subject Wood's theory to a detailed criticism. There are obvious defects in his exposition: for example, the implicit assumption that there are two homogeneous factors of production, labor and capital, and consequently such entities as "the" wage rate and "the" interest rate. The only point that seems worth elaborating here, however, is his peculiar concept of the nature of substitution. The essence of this concept is that production consists of integral processes which must be carried on exclusively by one of the two factors.[4] This is an unduly narrow view: Wood considers only what might be called the extensive margin of substitution between capital and labor, and overlooks the equally important possibilities of substitution within any process. But even a partial generalization of the substitution relationships among productive factors marked a major advance over the classical theory. And up to that time probably no other non-mathematical economist had shown equal insight into the interplay between production functions, demand conditions, and the supplies of productive factors.

2. "Theory of Wages," p. 21.
3. "New View," p. 468.
4. "Theory of Wages," p. 14: "The relative advantages of labor and of capital vary so greatly in their different uses, that whatever may be their hire at any time, it will in most of their employments be impossible to substitute one for the other, unless a vast change should occur in the relative price of using them." "New View," p. 69: " . . . at whatever point this equality of costs exists, labor and auxiliary capital will be employed indifferently; while on one side of it labor alone will be used and on the other side capital exclusively, . . . '

The History of Doctrine. Stuart Wood's first — and only voluntary — contribution to the history of ideas was a critical history of the wages-fund doctrine.[5] This historical survey was the most complete to appear down to that time, and it remains today an unusually illuminating and convincing appraisal.

The views of James Mill, McCulloch, Senior, and J. S. Mill are summarized and criticized because of their failure to face the problem of what determines the wages-fund. Torrens, and more particularly Cairnes, went farther: They assumed that the quantity of fixed capital and of materials (circulating capital not spent on labor) was determined by the size of the laboring population (since the proportion between capital and labor was fixed in each industry), and therefore the wages-fund is a residual.[6] Wood's criticism of this theory deserves full quotation:

> But, unfortunately, Cairnes, in his turn, fails to inquire into the causes of the existing industries and methods.
>
> One set of these causes is to be found in the character of the natural resources of any country, another in the state of the arts, and a third in the nature of human desires. If all of these things were invariable, or if, at least, none of them varied under the influence of price, then Cairnes's explanation would exhaust the causes which go to make up the rate of wages; . . . For then the national industries and their prevailing methods of production would be really fixed independently of the rate of wages, and would decide what stock of implements and materials must be assigned to each laborer. . . .
>
> But the price of labor does affect the choice of the industries and methods which prevail. It influences desire, because it affects the relative price of the objects of desire, in so far as labor enters into their production in different proportions. It therefore affects the choice between the industries which produce different kinds of goods, by reason of its influence on the desire for these goods. It also affects the choice between different methods of producing the same goods, by enhancing or cheapening the cost of those methods into which labor more largely enters, as compared with those into which the use of fixed capital enters more largely. In these ways the price of labor does potently affect the demand for labor, and no explanation of the phenomena of wages will hold good which ignores this fact.[7]

Thornton's refutation of the wages-fund theory by the argument that the demand curve for labor has zero elasticity is properly re-

5. "A Critique of Wages Theories," *Annals* of the American Academy of Political and Social Science, I (1890), 426–461.

6. Ibid., pp. 446, 453. Wood forgets the indeterminate deduction from total capital for the capitalists' consumption.

7. Ibid., pp. 452–453.

jected on essentially the same ground of substitution between factors.[8]

Only two points in Wood's appraisal of the wages-fund doctrine are seriously objectionable. The first is a positive error: he suggested that the classical economists believed that the size of the wages-fund was independent of labor's productivity,[9] when of course it was generally recognized that increases in productivity would lead to a larger future wages-fund. The second point is an oversight, in this as in his previous essays: Wood never questioned the realism or usefulness of the concept of an "average wage rate." But it cannot be disputed that he seized upon the greatest single defect in the classical theory of production and distribution, the failure to discover the relationship of substitution among productive factors.

Wood's final publication in economics was a review of Augusto Graziani's "Studii sulla Teorica Economica delle Macchine."[1] Graziani had criticized von Thünen's marginal productivity theory and extended the criticism to Wood's version, which had been presented "without too much novelty of reasoning."[2] Wood vehemently denied the accusation of plagiarism which he read into this passage, but the text supports Graziani's rejoinder that no insinuation of plagiarism was intended. Graziani's criticisms of the marginal productivity theory were wrong; the episode acquires some significance only because of a sequel.

In the September, 1894 issue of the Journal of Political Economy there appeared the anonymous note which is reproduced on page 648.[3] The note is too brief to permit of an unambiguous

8. Wood points out the terminological confusion whereby the classical economists and Thornton secured contradictory conclusions from the same premise, that the demand for labor was fixed (independent of its price). By demand for labor, the classical economists meant the quantity of commodities offered in exchange for labor, Thornton meant the quantity of labor demanded. Ibid., p. 460.

9. "No assiduity on the part of the laborers, no improvement in production, could raise wages, and just as surely, no degree of neglect or indolence could lower them" (ibid., p. 450).

1. Turin: Fratelli Bocca, 1891. The review was "Note on Professor Graziani's Economic Theory of Machinery," Annals of the American Academy of Political and Social Science, II (1891), 522-530; to which a reply was made by Graziani, "The Economic Theory of Machines," ibid., 838-841.

2. The disputed passage is on pp. 57-58 of Graziani's book, translated (with minor inaccuracies) in Wood's review, p. 526.

3. II (1894), 574. The present quotation follows the published note,

NOTES

An interesting coincidence between Lauderdale's "Public Wealth," published early in the century, and a recent study on wages, appears below, in parallel columns:

S. Wood, "New View of the Theory of Wages." QUARTERLY JOURNAL OF ECONOMICS, October, 1888, and July, 1889.

The price of a given amount of labor is equal to the price which is paid for the use of such amount of auxiliary capital as can replace it in those operations where the two things can be indifferently employed with equal pecuniary advantage (October, 1888, p. 68).

We have seen that capital and labor compete with each other for employment, being able often to supplant each other (p. 85).

If a man can be hired for one hundred and ten dollars to do a certain piece of work, and if, when interest is at six per cent, a machine can be hired for one hundred and twenty dollars that will do the same work, then so long as the rate of interest remains at six per cent labor will be employed, but if interest falls to five per cent then the machine will be employed (p. 64).

The absolute limit to the increase of capital would be attained when its amount should be at the maximum which can usefully be employed in connection with the existing supply of labor, in the existing state of nature and art (July, 1889, p. 478).

Lauderdale, "Public Wealth." [Edinburgh, 1804.]

That the profit of stock employed in machinery is paid out of a fund that would otherwise be destined to pay the wages of the labor it supplants is evident (p. 167) . . . He [the owner of machinery] has only to charge a little less than the wages of the labor which the machine supplants.

Supposing one man with a loom should be capable of making three pair of stockings a day, and that it should require six knitters to perform the same work with equal elegance in the same time, it is obvious that the proprietor of the loom might demand . . . the wages of five knitters, . . . but if a stocking loom was only capable of making one pair of stockings in three days . . . it would be thrown aside as useless (p. 165-6).

There must at all times be a point determined by the existing state of knowledge and the art of supplanting labor . . . beyond which capital cannot profitably be increased, . . . because . . . when it exceeds that point . . . its value must of consequence diminish in such a manner as effectually to check its augmentation (p. 228).

interpretation, but its wording and format are calculated to suggest plagiarism to most readers.' Wood never replied to the note.

We may reject the implication. Casual analogies are not a sufficient basis for so grave a charge. Brief study would reveal that Wood's theory differs in important respects from that of the eccentric Lord Lauderdale. The latter, it is true, emphasized the substitution of capital for labor. But he argued that only a monopolist can secure for his capital goods a price equal to the wages of the displaced labor;[4] under competition the return on machinery is set by "the proportion betwixt the quantity of machines that can be easily procured, and the demand for them." [5] Nor does Lauderdale see the possibility of substituting labor for capital. As the third set of parallel quotations in the note indicates, Lauderdale usually wrote as if the quantity of capital was governed by the state of the arts;[6] Wood placed this ultimate technological limitation on the use of capital in the background and focussed attention on the prices of the factors. But even if the theories of Wood and Lauderdale were similar in details, Wood's high character and his total lack of professional incentives would make plagiarism unthinkable.

GEORGE J. STIGLER.

COLUMBIA UNIVERSITY

which contains twenty-six minor errors. It has been impossible to determine the authorship of the note.

4. An Enquiry into the Nature and Origin of Public Wealth (Edinburgh, 1804), pp. 168–169, 179.

5. Ibid., p. 167; also p. 179.

6. Ibid., pp. 214–215, 219–221, 228; on p. 252, however, he inconsistently says that increases of capital may force capitalists to supplant labor "at a cheaper rate."

MULTIPLE-PLANT FIRMS: COMMENT

In his discussion of multiple-plant firms (this JOURNAL, February, 1947) Mr. Patinkin correctly states the well-known primary conditions of optimum capacity utilization at any given level of total output, which require that

(a) if two or more plants are operated simultaneously, the rates of output in all these plants must be such as to equate their marginal costs;

(b) no plant should be kept idle if its marginal costs at zero output are lower than the marginal costs of any other plant at its actual rate of operation.

He does not, however, state the important secondary conditions which also have to be satisfied if the given combined output of all plants is to be produced at lowest possible total costs, namely,

(c) not more than one plant should be operated at decreasing marginal costs;

(d) if one plant is actually operated at falling marginal costs, the reciprocal of their (negative) rate of decrease must not be smaller in its *absolute* magnitude than the reciprocals of the rates of increase of marginal costs of all other operating plants added together.

The equality of marginal costs of all operating plants (condition (a) above) as stated in mathematical terms — top of page 191 in his article — leads to radical simplification of the secondary conditions (inequalities) reproduced in their standard form on the same page below.

Subtracting from the first row of the smallest of his determinants, its second row multipled by c_1'/c_1'' and its third row multiplied by c_2'/c_2'' (I use c_i' for Mr. Patinkin's $\partial c_i/\partial x_i$ and c_i'' for his $\partial^2 c_i/\partial x^2{}_i$) one can reduce the first of his inequalities to the following simple expression: $(1/c_i''+1/c_2'') \cdot c_1'' \cdot c_2''=c_1'' + c_2'' \geq 0$.

The equality sign can be added to the (now reversed) inequality sign used by Mr. Patinkin if one is interested in all distributions of any given combined output between different plants which will minimize its total costs. Since the number "1" and "2" can be attached to any two different operating plants, the last expression states our rule (c).

All other inequalities written out on page 191 can be reduced to similar simple expressions. The last one in particular can be written as:

$$(1/c_1'' + 1/c_2'' + \cdots + 1/c_n'') \ c_1'' \cdot c_2'' \cdots c_n'' > 0.$$

If none of the second derivatives is negative, the condition will obviously be satisfied; if only one of them is negative, it will be satisfied, provided the expression in parenthesis is not positive. This situation is covered by rule (d) stated above. Rule (c) excludes all cases with more than one non-positive second derivative.

Mr. Patinkin's second and all other intermediate inequalities do not represent independent conditions, since they are necessarily fulfilled if conditions (c) and (d) are satisfied.

WASSILY W. LEONTIEF.

HARVARD UNIVERSITY

NOTE ON THE ALLOCATION OF OUTPUT[1]

Professor Leontief's comment provides two very interesting theorems on the allocation of output in the case of multiple-plant firms. His theorems hold for any output. Usually, however, we are concerned only with the output for which the firm maximizes profits; for this "optimum output" it is possible to derive stronger theorems than those developed by Professor Leontief.

Assume, for the sake of simplicity, that all plants are in operation. Let x_i be the output of the ith plant and $c_i = c_i(x_i)$ $(i = 1 \cdots n)$ be its total cost curve. Let the total revenue curve of the firm be $R = R\left(\sum\limits_{i=1}^{n} x_i\right)$. The firm then maximizes its profits

$$(1) \qquad R\left(\sum_{i=1}^{n} x_i\right) - \sum_{i=1}^{n} c_i(x_i)$$

with respect to the x_i to yield the first order necessary conditions for a maximum

$$(2) \qquad R' - c_i' = 0 \quad (i = 1 \cdots n)$$

These are the familiar conditions for the equality of marginal revenue and the marginal cost of each (operating) plant.

1. I am indebted to Kenneth J. Arrow (Cowles Commission, University of Chicago) for very helpful suggestions.

In general, the allocation of output among the plants derived from (2) will not be unique. Furthermore, some of these allocations may correspond to minimum profits. These are excluded by imposing the second order necessary conditions for a maximum:

$$
(3)\ (-1)^k
\begin{vmatrix}
R''-c_1'' & R'' & R'' & \cdots & R'' \\
R'' & R''-c_2'' & R'' & \cdots & R'' \\
R'' & R'' & R''-c_3'' & \cdots & R'' \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
\cdot & \cdot & \cdot & \cdots & \cdot \\
R'' & R'' & R'' & \cdots & R''-c_k''
\end{vmatrix}
\geq 0
\quad (k=1\cdots n)
$$

where R'' (c'') is the slope of the marginal revenue (cost) curve. For $k=1$ and $k=2$ this reduces to

$$(4) \qquad\qquad R'' \leq c_1''$$

and

$$(5) \qquad\qquad c_1''c_2'' \geq R''(c_1''+c_2'')$$

respectively. Now assume that both c_1'' and c_2'' are negative. Multiplying (4) by $(c_1''+c_2'')$ we obtain

$$(6) \qquad R''(c_1''+c_2'') \geq c_1''(c_1''+c_2'') = (c_1'')^2 + c_1''c_2''.$$

Comparing (5) with (6) we see that (assuming $c_1''\neq 0$) there is a contradiction. Since any two plants can be considered as the "first and second," this proves that at the "optimum output" not more than one plant will be operating on the falling part of its marginal cost curve. This is Professor Leontief's statement (c) applied to a specific output: the optimum market output.

By this result we can eliminate some of the allocations determined by (2). The remaining admissible ones will all be of one of the following types: (a) all plants operating at the same marginal cost on rising parts of their respective marginal cost curves; and (b) all plants operating at the same marginal cost, but one operating on the falling part of its curve, and the rest on the rising part of theirs. However, we have used only the first two conditions of (3). By using the n-th, even some of these remaining allocations

can be rejected as not satisfying the second order necessary conditions for a maximum..

Consider now (3) for $k=n$. Subtract the last row from each of the other rows. This reduces (3) to

$$(7) \quad (-1)^n \begin{vmatrix} -c_1'' & 0 & \cdots & 0 & c_n'' \\ 0 & -c_2'' & \cdots & 0 & c_n'' \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & -c_{n-1}'' & c_n'' \\ R'' & R'' & \cdots & R'' & R''-c_n'' \end{vmatrix} \geqq 0$$

Now multiply the j-th row ($j=1 \cdots n-1$) by $\dfrac{R''}{c_j''}$ and add to the last one, so that (7) reduces to

$$(8) \quad (-1)^n \begin{vmatrix} -c_1'' & 0 & \cdots & 0 & c_n'' \\ 0 & -c_2'' & \cdots & 0 & c_n'' \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ \cdot & \cdot & \cdots & \cdot & \cdot \\ 0 & 0 & \cdots & 0 & R''-c_n''+R''c_n''\left(\sum\limits_{j=1}^{n-1}\dfrac{1}{c_j''}\right) \end{vmatrix}$$

$$= (-1)^n(-1)^{n-1}\left[\prod_{j=1}^{n-1}c_j''\left\{R''-c_n''+R''c_n''\left(\sum\limits_{j=1}^{n-1}\dfrac{1}{c_j''}\right)\right\}\right]$$

$$\geqq 0.$$

Multiplying and dividing by c_n'' we have

$$(9) \qquad \prod_{i=1}^{n} c_i''\left(R''\sum\limits_{i=1}^{n}\dfrac{1}{c_i''}-1\right) \leqq 0.$$

There are several cases to consider: (α) allocation (a) with $R''<0$; (β) allocation (a) with $R''>0$; (γ) allocations (b) with $R''<0$; and (δ) allocation (a) with $R''=0$. Note that by (4) (since the plant operating on the declining part of its marginal cost curve can be considered as the "first") there are no cases of allocation (b) and $R'' \geqq 0$.

(α) Since $\Pi c_i''>0$ and $R''<0$, (9) reduces to

$$(10) \qquad \sum_{i=1}^{n} \frac{1}{c_i''} \geq \frac{1}{R''}$$

By assumption, the left side of this inequality is positive, and the right negative; therefore it is always satisfied. Consequently, it is impossible to reject (on the basis of (9)) an allocation (a) when $R'' < 0$.

(β) Since $\Pi c_i'' > 0$ and $R'' > 0$, (9) reduces to

$$(11) \qquad \sum_{i=1}^{n} \frac{1}{c_i''} < \frac{1}{R''}$$

where, by assumption, both sides of the inequality are positive. In this case it is possible to reject allocation (a) if the c_i'' are not sufficiently large.

(γ) Since $\Pi c_i'' < 0$ and $R'' < 0$, this reduces to

$$(12) \qquad \sum_{i=1}^{n} \frac{1}{c_i''} < \frac{1}{R''}.$$

This will not be satisfied if the absolute value of the reciprocal of the negative c_i'', does not exceed the sum of the reciprocals of the marginal costs of the other plants by an amount $\geq$ | Thus some of the allocations (b) can be rejected on the basis of (12).

(δ) Since $R'' = 0$, (9) reduces to

$$(13) \qquad -\Pi c_i'' \leq 0.$$

This is always satisfied for any allocation (a).

These results can be summarized in the following

Theorem: All the allocations that will maximize profits are included under types (a) or (b). Allocation (a) can be rejected in the case where $R'' > 0$ and (11) is not satisfied. Allocation (b) can be rejected for $R'' < 0$, if (12) is not satisfied; and for $R'' \geq 0$, in all circumstances.

As was to be expected, the conditions of rejection formulated in this theorem are broader than those proposed by Professor Leontief. This is due to the fact that he discusses all outputs, while I am concerned only with the optimum output. Thus, Professor Leontief does not reject any allocation (a) (c.f. first sentence of next to last paragraph of his note), while our theorem does reject some of these in the case $R'' > 0$ (cf. (β)). Similarly, for allocations (b), Professor Leontief rejects if "the reciprocal of

the (negative) rate of decrease is greater in its *absolute* magnitude than the reciprocals of the rates of increase of marginal costs of all other plants added together." In our theorem, we reject (in the case of $R''<0$) if the absolute magnitude does not exceed this sum by an amount $\geqq \left| \dfrac{1}{R''} \right|$.

The reader can establish for himself the fact that there are only three independent restrictions in (3): $k=1, 2, n$. If these are satisfied, those for $k=3 \cdots n-1$ are also satisfied.

The fundamental point here is that even on the basis of our theorem, there may still remain several unrejected allocations. Furthermore, it is possible that they will satisfy the sufficient conditions for a relative maximum, viz., (3) with the equality sign removed. Under these assumptions, each of the remaining allocations will then maximize profits relative to all allocations within a neighborhood. But we are interested in the *absolute* maximum — i.e. the highest of all these relative maxima. (This is analogous to the discussion in my article under point 3, page 189 (this Journal, February, 1947)). For this purpose, conditions (3) can give us no additional information. It may be necessary actually to compute the profit at these different allocations in order to discover the one which gives the highest profit.

There is an interesting special case of the preceding analysis. This is the case of perfect competition. By this assumption $R''=0$ so (9) reduces to

$$(14) \qquad\qquad -\prod_{i=1}^{n} c_i'' \leqq 0.$$

This establishes the following

Corollary: Under perfect competition, it is impossible to have any plant of a multiple-plant firm operating on the declining part of its marginal cost curve.

So far, we have abstracted from the possibility of non-operating plants. Introducing this possibility complicates the problem considerably. Let us consider a fixed output x, and determine the conditions under which a plant will not be operated. These have been stated by Professor Leontief under his propositions (*a*) and (*b*); but it is interesting to consider the rigorous analytic proof of these propositions.

The firm minimizes its costs $\sum\limits_{i=1}^{n} c_i(x_i)$ subject to the restraint that

$$(15) \qquad \sum_{i=1}^{n} x_i = x > 0.$$

For our purposes we must specifically introduce the additional restrictions, always tactitly assumed,

$$(16) \qquad x_i \geq 0 \qquad i = 1 \cdots n$$

i.e. no plant can produce a negative output. It can be shown that the problem of minimizing total costs subject to (15) and (16), is equivalent to minimizing

$$(17) \qquad \sum_{i=1}^{n} c_i(x_i) - \lambda \left(\sum_{i=1}^{n} x_i - x \right) - \sum_{i=1}^{n} \mu_i x_i$$

with respect to the x_i to yield the n equations

$$(18) \qquad c_i' - \lambda - \mu_i = 0 \qquad (i = 1 \cdots n)$$

(where λ and μ_i are Lagrange multipliers), and imposing the $n+1$ restrictions

$$(19) \qquad \mu_i x_i = 0 \qquad (i = 1 \cdots n)$$

and (15). By (15), at least one of the x_i is greater than zero. Let this be x_1. Then by (19), $\mu_1 = 0$. Substituting in (18) we obtain

$$(20) \qquad c_1' = c_j' - \mu_j \qquad j = 2 \cdots n$$

There are two possibilities (a) $x_j > 0$; (b) $x_j = 0$. If $x_j > 0$, then by (19) $\mu_j = 0$. Substituting into (20), we see that the marginal cost of all operating plants must be equal. Now assume $x_j = 0$. Then there are two possibilities to consider: $\mu_j < 0$ and $\mu_j \geq 0$. Assume $\mu_j < 0$; then by (20)

$$(21) \qquad c_1'(x_1) > c_j'(0)$$

i.e. the marginal cost of the operating plant is greater than that of the idle one. Consider the expression

$$(22) \qquad c_1(x_1 - t) + c_j(t)$$

as a function of t, where t is restricted to positive values. Differentiating this function at $t = 0$, and employing (21) we obtain, as the expression for the derivatives

$$(23) \qquad -c_1'(x_1) + c_j'(0) < 0$$

Hence, by the continuity of $c_1(x_1)$ and $c_j(x_j)$, it follows from (23) that for sufficiently small (positive) t

$$(24) \qquad c_1(x_1 - t) + c_j(t) < c_1(x_1) + c_j(0).$$

The left member of (24) represents total cost when the first plant produces $x_1 - t$, and the j-th, t, The right member represents total cost for the allocation x_1, 0. Hence assuming $\mu_j < 0$ leads to a contradiction of the assumption that the allocation x_1, 0 minimizes costs. Thus if $x_j = 0$, we must have $\mu_j \geq 0$. Consulting (20), we see that this proves that every non-operating plant will have marginal costs (at zero output) greater than the marginal costs of any operating plant (at its actual rate of output). It can be readily proved that this same rule holds (without further restriction) for the profit-maximizing output too.

There are many additional (and very interesting) problems with respect to the allocation of output in multiple-plant firms. First of all, it may be possible to set an upper bound to the number of admissible allocations corresponding to relative profit-maximizing outputs. Second, 1 conjecture that it is possible to have situations where outputs of $x = x_0$ are produced with j-plants, $x = x_0 + d$ with $j+1$ plants, and $x = x_0 + d + e$ with j plants again. I have not yet been able to develop the conditions under which this would be true. Finally, there is the very general problem of choosing among relative maximum or minimum positions. I hope to treat some of these problems in the future.

I should like to take advantage of this opportunity to correct the following errors which unfortunately crept into my article: On page 183, line 9 from bottom, replace "s_j and s_{j+1}" by "c_j and c_{j+1}." On page 185, last line; page 186, line 21; page 189, lines 11 and 20 — replace "equalization of marginal cost" by "equalization of marginal cost by equalizing outputs."

<div align="right">DON PATINKIN.</div>

UNIVERSITY OF CHICAGO

RECENT PUBLICATIONS

ALLAIS, M. Abondance ou Misère. Paris: Librairie de Medicis. 1946. pp. 131.

BAIN, JOE S. The Economics of the Pacific Coast Petroleum Industry. Berkeley and Los Angeles, California: The University of California Press. 1947. pp. 123. $3.00.

BAKER, HELEN, and TRUE, JOHN M. The Operation of Job Evaluation Plans. Princeton, New Jersey: Industrial Relations Section, Department of Economics and Social Institutions, Princeton University. 1947. pp. 111. $1.50.

BENHAM, FREDERIC, and BODDY, FRANCIS M. Principles of Economics. New York: Pitman Publishing Corporation. 1947. pp. 430. $3.50.

BROWN, ALVIN. Organization of Industry. New York. Prentle-Hall, Inc. 1947. pp. 370. $4.00.

BROWN, ESTHER LUCILE. The Use of Research by Professional Associations in Determining Program and Policy. New York: Russell Sage Foundation. 1946. pp. 39. 25 cents.

BROWN, LEO CYRIL, S.J. Union Policies in the Leather Industry. Cambridge, Massachusetts: Harvard University Press. 1947. pp. 246. $3.50.

BURNS, ARTHUR F. Stepping Stones towards the Future. New York: National Bureau of Economic Research. 1947. pp. 91. [Twenty-Seventh Annual Report of the Bureau.]

CASSELMAN, PAUL HUBERT. Methods Employed in Estimating Total Manpower on War Work in Canada. Ontario, Canada: New Era Publications. 1947. pp. 12. 15 cents. [Technical Series Publications No. 2.]

CHURCHMAN, C. WEST, ACKOFF, RUSSELL L., WAX, MURRAY. Measurement of Consumer Interest. Philadelphia: University of Pennsylvania Press. 1947. pp. 214. $3.50.

CIVILIAN PRODUCTION ADMINISTRATION, BUREAU OF DEMOBILIZATION. Field Organization and Administration of the War Production Board and Predecessor Agencies May 1940 to November 1945. Washington, D. C.: Historical Reports on War Administration: War Production Board. 1947. pp. 288. [Special Study No. 25.]

COALE, ANSLEY J. The Problem of Reducing Vulnerability to Atomic Bombs. Princeton, New Jersey: Princeton University Press. 1947. pp. 116. $2.00.

CONDLIFFE, J. B. Obstacles to Multilateral Trade. Washington, D. C.: National Planning Association. 1947. pp. 46. 25 cents.

DAVIS, HIRAM S. Economic Progress. Philadelphia: University of Pennsylvania Press. 1947. pp. 187. $2.75.

DE MOND, ALBERT LAWRENCE. Certain Aspects of the Economic Development of the American Negro, 1865–1900. Washington, D. C.: Catholic University of America Press. 1945. pp. 187.

DUPRIEZ, LEON H. Monetary Reconstruction in Belgium. New York: King's Crown Press. 1947. pp. 88. $2.25.

ENKE, S., and SALERA, V. International Economics. New York: Prentice-Hall, Inc. 1947. pp. 731. $4.00.

GAMBINO, AMEDEO. Gli Ammodernamenti della Teoria Degli Scambi Internazionali. Milan: Università Commerciale Luigi Bocconi. 1946. pp. 60.

GINSBURG, DAVID. The Future of German Reparations. Washington, D. C.: National Planning Association. 1947. pp. 79. 50 cents. [Planning Pamphlet Nos. 57–58.]

HICKMAN, C. ADDISON. World Economic Problems. New York: Pitman Publishing Corporation. 1946. pp. 400.

JACOBY, NEIL H., and SAULNIER, RAYMOND J. Business Finance and Banking. New York: National Bureau of Economic Research. 1947. pp. 241. $3.50.

JUNG, NAWAB MIR NAWAZ and IYENGAR, S. KESAVA. The A, B, C of Central Banking. Bangalore City, India: The Bangalore Printing and Publishing Co., Ltd. 1947. pp. 552. 2000 *rupees*.

KÄFER, DR. KARL. Johann Friedrich Schär und die Entwicklung der Betriebswirtschaftslehre. Zurich, Switzerland: Schulthess & Company. 1946. pp. 13.

KIRZ, MIROSLAV A. Postwar International Lending. Princeton, New Jersey: International Finance Section, Department of Economics, Princeton University. 1947. pp. 29. [Essays in International Finance, No. 8.]

KUZNETS, SIMON. National Product since 1869. New York: National Bureau of Economic Research, Inc. 1947.

LANDIS, PAUL H. Social Policies in the Making. Boston. D. C. Heath & Company. 1947. pp. 554. $4.00.

LANGER, WILLIAM. Our Vichy Gamble. New York: Alfred A. Knopf. 1947. pp. 412. $3.75.

LAVIN, JOSE DOMINGO. Plan Inmediato de Industrializacion en Mexico. Mexico, D. F.: The Author. 1945. pp. 18.

LIBRARY OF CONGRESS. Annual Report, 1946. Washington, D. C.: United States Government Printing Office. 1947. pp. 538.

LOCKLIN, D. PHILIP. Economics of Transportation. Chicago: Richard D. Irwin, Inc. 1947. pp. 885. $5.50.

LURIE, SAMUEL, Private Investment in a Controlled Economy. New York: Columbia University Press. 1947. pp. 243. $3.00.

MARLIO, LOUIS. The Aluminum Cartel. Washington, D. C.: The Brookings Institution. 1947. pp. 130. $1.50.

MODLIN, GEORGE MATTHEWS amd DE VYVER, FRANK T. Development of Economic Society, Volume I. Boston: D. C. Heath & Company. 1946. pp. 474.

MUSSEY, HENRY R., and DONNAN, ELIZABETH. Economic Principles and Modern Practice. Boston: Ginn & Company. 1947. pp. 834. $4.50. [Second Edition.]

MYRDAL, GUNNAR. The Reconstruction of World Trade and Swedish Trade Policy. Stockholm: Svenska Handelsbanken. 1946. pp. 30.

NATIONAL ASSOCIATION OF MANUFACTURERS, ECONOMIC PRINCIPLES COMMISSION. The American Individual Enterprise System. Volumes I and II. New York: The Commission. 1946. pp. 1119. $10.00.

NATIONAL COMMITTEE ON IMMIGRATION POLICY. Economic Aspects of Immigration. New York: The Committee. 1947. pp. 63.

NATIONAL INDUSTRIAL CONFERENCE BOARD. What if Prices were Free? New York: The Board. 1946. pp. 43. 50 cents.

——. Will the Guaranteed Annual Wage Work? New York: The Conference Board. 1946. pp. 44. 50 cents.

NEUMANN, JOHN VON, and MORGENSTERN, OSKAR. Theory of Games and Economic Behavior. Princeton, New Jersey: Princeton University Press. 1947. pp. 641. $10.00.

OXFORD UNIVERSITY INSTITUTE OF STATISTICS. Studies in War Economics. Oxford, England: Basil Blackwell. 1947. pp. 410. 25s.

PALMER, GLADYS L. Research Planning Memorandum on Labor Mobility. New York: Social Science Research Council. 1947. pp. 22. 25 cents.

PLATIER, A. Le Contrôle des Changes. Paris: Institut International de Coöperation Intellectuelle. 1947. pp. 210.

PRAY, JOSEPH. Post-Appropriation Budgetary Control in Oklahoma. Norman, Oklahoma: The Author. 1946. pp. 27. 50 cents.

RESERVE BANK OF INDIA. Reserve Bank of India Bulletin. Bombay, India: The Bank. 1947. pp. 72.

SHEPHERD, GEOFFREY S. Marketing Farm Products. Ames, Iowa: Iowa State College Press. 1946. pp. 445. $4.25.

SLICHTER, SUMNER H. The Challenge of Industrial Relations. Ithaca, New York: Cornell University Press. 1947. pp. 196. $2.50.

SOCIAL SCIENCE RESEARCH COUNCIL. Labor Force Definition and Measurement. New York: The Council. 1947. pp. 134. $1.00.

STOCKTON, JOHN R. An Introduction to Business Statistics. Boston: D. C. Heath & Company. 1947. pp. 478. $4.00. [Second edition.]

SURVEL, JAAK. East and West — Facts from behind the Iron Curtain. London: The Baltic Review. 1947. pp. 70.

TARSHIS, LORIE. The Elements of Economics. Boston: Houghton Mifflin Company. 1947. pp. 699. $4.50.

TAX INSTITUTE SYMPOSIUM. How Should Corporations be Taxed? New York: Tax Institute. 1947. pp. 251. $4.00.

TOWLE, LAWRENCE W. International Trade and Commercial Policy. New York: Harper & Brothers. 1947. pp. 780. $4.50.

TROXEL, EMERY. Economics of Public Utilities. New York: Rinehart & Company, Inc. 1947. pp. 892. $5.75.

UNITED STATES DEPARTMENT OF COMMERCE, BUREAU OF THE CENSUS. A Chapter in Population Sampling. Washington, D. C.: The Department. 1947. pp. 141. $1.00.

UNITED STATES STRATEGIC BOMBING SURVEY. The Effects of Strategic Bombing on Japan's War Economy, Appendix A, B, C. Washington, D. C.: Over-All Economic Effects Division. 1946. pp. 244. 65 cents.

VICKREY, WILLIAM. Agenda for Progressive Taxation. New York: Ronald Press Company. 1947. pp. 496. $4.75.

WALDRON, GLORIA, and BUCHANAN, NORMAN S. America's Stake in World Trade. New York: Twentieth Century Fund. 1947. pp. 32. 20 cents.